Etwas Werbung:

Im Webshop unter http://shop.wikipedia.org findet sich eine Druckausgabe des WikiReaders zum Preis von 5,20 Eur.

Auch der WikiReader Schweden wurde gedruckt und ist dort erhältlich.

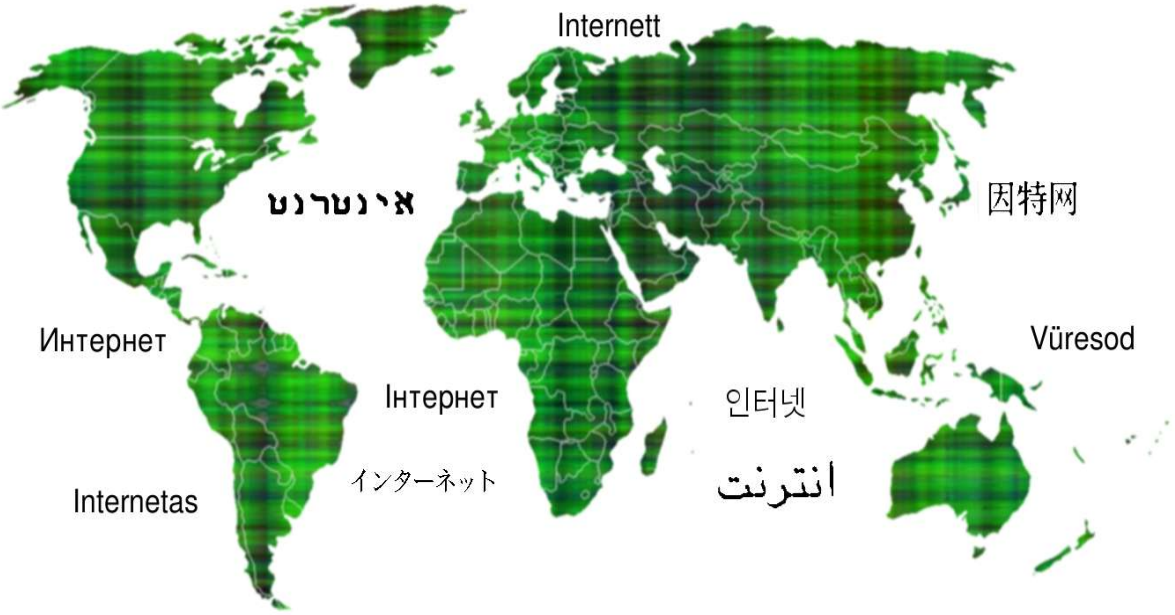
Für den WR Internet einfach weiterblättern.

WikiReader

Eine Artikelsammlung aus Wikipedia, der freien Enzyklopädie

Stand vom 13. Mai 2004

Internet



WIKIMEDIA FOUNDATION

WIKIPEDIA
Die freie Enzyklopädie

Impressum

Verfasser: Die Freiwilligen Schreiber der deutschsprachigen Wikipedia

Herausgeber dieser Ausgabe: , Thomas R. „TomK32“ Koll
Besonders fleißige Wikipedianer: chriss84, Erik Moeller, JeLuF, gwicke, Harko, Head, Hella, Martinroell, mwka, presroi, Sina Eetezadi, Tkarcher

Verwendete Schriften: FreeSerif und FreeMono
Titelbild: Eigenkreation von TomK32, Karte aus der Wikipedia

Stand dieser Ausgabe: 13. Mai 2004 um 13:40 MESZ

Adresse der Wikipedia: http://de.wikipedia.org
Adresse dieses Hefts: http://de.wikipedia.org/wiki/Wikipedia:WikiReader/Internet

ISSN (Onlineausgabe): 1613-7752
ISSN (Druckausgabe): noch nicht bekannt

Eine vollständige Liste der verwendeten Artikel und die Namen der angemeldeten Benutzer die an diesen mitgewirkt haben findet sich im Anhang des WikiReader

Über Wikipedia

Die Wikipedia ist eine freie Enzyklopädie, die es sich zur Aufgabe gemacht hat, jedem eine freie Wissensquelle zu bieten, an der er nicht nur passiv durch lesen teilhaben kann, sondern auch aktiv als Autor mitwirken kann. Auf der Webseite http://de.wikipedia.org findet man nicht nur die aktuellen Artikel der deutschsprachigen Wikipedia, sondern darf auch sofort und ohne eine Anmeldung mit schreiben. Auf diese Art sind seit 2001 in nur drei Jahren eine halbe Million Artikel zustande gekommen, in mehr als 40 Sprachen.

Inzwischen ist die Wikipedia seit 2003 Teil der Wikimedia Foundation die für die technischen Voraussetzungen der Wikipedia zuständig ist und auch andere Projekte wie das Wörterbuch Wiktionary oder das Lehrbuch-Projekt WikiBooks beherbergt.

Über die Reihe WikiReader

WikiReader ist eine unregelmäßig erscheinende Heftreihe, welche ausgewählte Wikipedia-Artikel thematisch bündelt und in einer redaktionell aufbereiteten Form präsentiert. Die Auswahl der Artikel erhebt keinen Anspruch auf Vollständigkeit oder Korrektheit, sondern soll gewissermaßen als „Schnappschuss“ des jeweiligen Themas dienen. Wir ermuntern unsere Leser ausdrücklich dazu, selbst weiter zu recherchieren, Artikel in der Wikipedia zu verbessern oder auch neue Artikel hinzuzufügen und damit Anregungen für zukünftige WikiReader-Ausgaben zu liefern.

Über den WikiReader Internet

Die vorliegende Ausgabe der Heftreihe hat es sich zum Ziel gesteckt, die vielfältigen Aspekte des Internet zu beleuchten und dem interessierten Laien eine fundierte Wissensgrundlage zu bieten.

Nachdem der erste WikiReader von Thomas Karcher zum Thema Schweden im Feber 2004 veröffentlicht wurde, kamen schnell Vorschläge für neue WikiReader auf. Von den vorgeschlagenen Themen wies Internet die meisten Artikel auf. Der vorliegende Reader war kaum mit dem Schreiben neuer Artikel verbunden, sondern ist fast ausschließlich eine Zusammenstellung aus bestehenden Artikel die aber dennoch auf einem sehr aktuellen Stand sind. Anfangs war kaum ein Artikel älter als drei Monate, ein gutes Zeichen für die Aktualität der Wikipedia. Inzwischen wurde dies sogar noch verbessert auf unter einen Monat.

Rechtliches

Inhaltsverzeichnis

Impressum....1
Über Wikipedia....2
Über die Reihe WikiReader....2
Über den WikiReader Internet....2
Rechtliches....2

Die Anfänge des Internet....6
ARPANET....6
Internet....6

Technik des Internet....11
Computernetzwerk....11
Topologien....11
Drahtlose Netze....13
Netzwerkprotokoll....14
OSI-Modell....14
Die Protokollschichten....15
Internet-Protokoll-Familie....17
Internet Protocol....20
IPv4....20
Routing....21
ICMP....22
Vergangenheit und Zukunft....24
IPv6....24
Effizienzsteigerungen....27
IPv6-Verfügbarkeit....28
UDP....30
Transmission Control Protocol....31
Zuverlässigkeit....32
HTTP....33
HTTP-Statuscodes....35
FTP....36
ftp-Kommandos....37
Telnet....38
Sicherheit....39
Secure Shell....39
Domain Name System....40
Komponenten des DNS....41
Resource Records....42
Nameserver....43
Resolver....44
DynDNS....45
DNS-Sicherheit....45
Domain-Registrierung....45
Domain....45
Top Level Domain....46
Länderspezifische Top Level Domains....47
Subdomain....51
Root-Server....51

Verbindung zum Internet....53
Akustikkoppler....53
Modem....53
ISDN....54
Digital Subscriber Line (DSL)....57
Point-to-Point Protocol....59
PPP over Ethernet....60
Standleitung....60
Wireless LAN....61
Sicherheit....63
Internetdienstanbieter / Provider....64
Client-Server-System....64
Network Address Translation....65
Proxy....66
Routing....68
Ping....70
Traceroute....71
Download....71
Upload....72

Dienste im Internet....73
Mailbox....73
E-Mail....74
Klammeraffe....76
Mailingliste....77
vCard....77
GNU Privacy Guard....78

Gopher....79
Usenet....80
Newsgroups....81
TOFU....84
Webfoum....85
Internet Relay Chat....86
Chat....87
Instant Messaging....87
IP-Telefonie....88
Vorteile der IP-Telefonie....90
Entwicklung....91
Peer-to-Peer....92
Napster....95
BitTorrent....96
Suchmaschine....97
World Wide Web....98
Webseite....100
Hypertext....101
Hyperlink....102
Uniform Resource Identifier....102
Webbrowser....104
Mozilla....105
Bookmark / Lesezeichen....107
Webdesign....108
Auszeichnungssprache....109
HTML....110
Falsche Interpretation von Webdokumenten....113
Weiterentwicklungen....114
HTML lernen....115
XML....115
Aufbau einer XML-Datei....116
Java-Applet....118
Cookie....119
Common Gateway Interface (CGI)....120
PHP....121
Webserver....123
LAMP....124
Barrierefreies Internet....124
Deutschland....125
Weblog....126
Wiki....127
Webportal....128

Andere Dienste und Verbindungsarten....131
WAP....131
I-mode....132
UMTS....132
Internet2....134

Sicherheit im Internet....135
Computersicherheit....135
Firewall....136
DMZ....137
IPsec....138
Kritik an IPsec....141
Security through Obscurity....142

Gefahren im Internet....144
Cracker....144
Denial of Service....145
IP-Spoofing....147
Computervirus....148
Geschichte....150
Prävention....150
Computerwurm....151
Trojanisches Pferd....153
Backdoor....154
Rootkit....154
Dialer....155
Gesetzliche Regelungen....157
Spyware....157
Web-Bug....158
Spam....158
Gesetze....160
E-Mail-Filter....161

Wichtige Organisationen....165
ICANN....165
Internet Engineering Task Force....165
Internet Research Task Force....166
RFC....167
World Wide Web Consortium....168

Internetarchiv.....168
DENIC.....168
Chaos Computer Club.....169

PERÖNLICHKEITEN.....171

Jonathan Postel.....171
Robert E. Kahn.....171
Vinton Gray Cerf.....172
Eric Allman.....173
Tim Berners-Lee.....174
Al Gore und das Internet.....175

MENSCH UND INTERNET.....177

Netzkultur.....177
Internetsucht.....177
Emoticon.....179
Englische Sprache im Internet.....180
Netiquette.....181
Troll.....182
E-Business.....183
E-Government.....185
Zensur im Internet.....187

APPENDIX.....189

Autoren.....189
Quellenverzeichnis.....190
GNU Freie Dokumentationen Lizenz .192

Die Anfänge des Internet

ARPANET

Das **ARPANET** wurde ursprünglich im Auftrag der US-Luftwaffe in Erwartung eines Atomkriegs ab 1962 von einer kleinen Forschergruppe unter der Leitung von Paul Baran entwickelt. Es ist der Vorläufer des heutigen Internet.

Es sollte ein dezentrales Netzwerk geschaffen werden, so dass im Kriegsfall die Kommunikation auch bei Ausfall vieler Knotenpunkte weiter möglich gewesen wäre. Das damals revolutionäre Konzept enthielt schon die grundlegenden Aspekte des heutigen Internet. Die Verbindungen wurden über Telefonleitungen hergestellt.

Das Projekt wurde zunächst vom Pentagon abgelehnt, im Jahre 1965 jedoch wieder aufgegriffen und 1969 realisiert. Anfangs vernetzte das Netzwerk lediglich die vier Forschungseinrichtungen *Stanford Research Institute*, *University of Utah*, *University of California in Los Angeles* und die *University of California in Santa Barbara*.

Zur selben Zeit wurde das Betriebssystem UNIX und die Programmiersprache C entwickelt. Diese 3 Komponenten entstanden unabhängig voneinander - doch die Zusammenführung von C, Unix und dem Arpanet trug wesentlich zur Entstehung des heutigen Internet bei. UNIX wurde in der Programmiersprache C umgeschrieben und war so auf vielen Maschinenplatformen verfügbar und erweiterbar, das erleichterte die Entwicklung von Kommunikationsanwendungen und Protokollen erheblich. Das Arpanet sorgte für eine einheitliche Möglichkeit, über weite Strecken zu kommunizieren, so wie es heute alltäglich ist.

Internet

Das Internet ist ein weltweites Computernetzwerk. Es dient der elektronischen Kommunikation und dem Austausch von Informationen.

Das Internet ging Ende der 1960er Jahre aus dem militärischen ARPANET hervor, einem Projekt der ARPA. Es wurde später benutzt, um Universitäten und Forschungseinrichtungen zu vernetzen, zunächst in den USA, später dann auch weltweit.

- 1969 die Network Working Group wird gegründet und erstellt die ersten Protokollbeschreibungen; die ersten vier Knoten des ARPANETs gehen in Betrieb
- 1971 das ARPANET besitzt 15 Knoten. Telnet und ftp werden entwickelt.
- 1972 **Ray Tomlinson** entwickelt das erste E-Mail-Programm
- 1973 das Transmission Control Protocol (TCP) wird publiziert
- 1977 das ARPANET besitzt 111 Knoten

- 1982 das spätere EUnet-Projekt der Informatik-Rechner-Betriebsgruppe (IRB) (Fachbereich Informatik, Universität Dortmund) unter Leitung von Dr. Rudolf Peter bietet erste Netzwerkdienste in Deutschland an.
- 1983 das ARPANET hat 400 angeschlossene Rechner
- 1984 das Domain Name System (DNS) wird entwickelt. Das ARPANET hat 1.000 angeschlossene Rechner.
- 1987 der Begriff "Internet" entsteht, es sind nun 27.000 Rechner vernetzt
- März 1989 **Tim Berners-Lee** schreibt die erste Fassung seines *„Information Management: A Proposal“*, der erste "Entwurf" für das WWW. Siehe auch *„A Little History of the World Wide Web“*
- Anfang 1989 erste **deutsche Internetanschlüsse** (Projekt EUnet (siehe 1982), Universität Dortmund, Dr. Rudolf Peter; Arbeitsgruppe Xlink, Prof. Zorn, Universität Karlsruhe); Details finden sich unter http://www.netplanet.org/geschichte/deutschland.shtml
- 1990 das militärische ARPANET wird außer Betrieb genommen
- November 1990 **Tim Berners-Lee** und **Robert Cailliau** veröffentlichen das Konzept für ein weltweites Hypertext-Projekt
- 1991 das WWW wird am CERN eingesetzt
- Dezember 1992 wird das ehemalige Forschungsprojekt und Netzwerkvorreiter in Deutschland **EUnet** wird privatisiert. Mit der EUnet Deutschland GmbH entsteht der erste kommerzielle Internet-Provider Deutschlands in Dortmund. Der Slogan: *"Connecting Europe since 1982"*. EUnet wird später an UUnet verkauft.
- 1993 WWW-Software wird außerhalb des CERN eingesetzt.
- Mai 1993 die Informatik-Rechner-Betriebsgruppe (IRB) (Fachbereich Informatik, Universität Dortmund) richtet mit ihrem Webauftritt einen der ersten öffentlichen Web-Server in Deutschland ein. Deutschlandweit gibt es zu dieser Zeit weniger als 15 Webserver.
- 1993 ein Jahr nach EUnet wird auch **XLink** privatisiert. Es entsteht damit der zweite Internet-Provider in Deutschland.
- August 1993 im August Gründung der IV-**DENIC** als zentraler Registrar für .de-Domains
- Oktober 1993 es gibt ca. 500 WWW-Server weltweit
- 1994 die Zahl der kommerziellen Nutzer im Internet übersteigt erstmals die der wissenschaftlichen Nutzer. Es gibt ca. 3 Millionen Internet-Rechner.
- 1997 das Projekt *Abilene* für ein Internet2 wird gestartet
- Oktober 1998 die ICANN wird gegründet
- Oktober 1999 die einmillionste .de-Domain wird registriert
- Seit dem 1. März 2004 sind auch **Umlaute und Sonderzeichen** in .de- und .ch-Domains erlaubt
- Im März 2004 sind bei der Denic über sieben Millionen .de-Domains registriert

Die anfängliche Entstehung und Verbreitung des Internet ist eng mit der Entwicklung des Betriebssystem UNIX verbunden.

Starken Auftrieb erhielt das Internet seit Anfang der 1990er durch das World Wide Web, kurz WWW. Mit Webbrowsern konnten nun auch Laien auf das Netz zugreifen. Durch die wachsende Zahl von Nutzern wurden auch viele kommerzielle Angebote ins Netz gestellt. Das Internet ist ein wesentlicher Katalysator der **Digitalen Revolution**.

Jahr	Hosts
1981	200
1983	500
1985	2.000
1990	313.000
1995	6.600.000
2000	93.000.000
2003	172.000.000

Anzahl der Rechner (gerundet) im Internet
(*http://www.isc.org/ds/host-count-history.html*)

Neue Trends im Internet verändern das Netz und ziehen neue Benutzerkreise an: IP-Telefonie, Kollaborationssoftware (Groupware) wie Wikis, Breitbandzugänge (z. B. für Video on Demand) und Peer2Peer-Vernetzung, (vor allem für Tauschbörsen).

Technik

Das Internet basiert auf der einheitlichen TCP/IP-Protokollfamilie, die einen Standard für Adressierung und Datenaustausch zwischen verschiedenen Computern und Netzwerken festlegt. Daraus ergibt sich der Name Inter-Network (Inter = lat. zwischen) - also das (Über-)Netzwerk, das die vereinzelten Netzwerke miteinander verbindet. Ein großer Vorteil ist es, dass die Kommunikation völlig unabhängig von den verwendeten Betriebssystemen und Netzwerktechnologien geschehen kann.

Das Domain Name System, abgekürzt DNS, ist ein wichtiger Teil der Internet-Infrastruktur. Um einen bestimmten Computer ansprechen zu können, identifiziert ihn das IP-Protokoll mit einer eindeutigen IP-Adresse. Dabei handelt es sich bei der heute üblichen Version IPv4 um 4 Byte (Zahlen im Bereich von 0 bis 255), die durch einen Punkt getrennt angegeben werden (z. B. 214.235.81.190). Man kann sich diese Zahl als eine Art Telefonnummer mit dem DNS als Telefonbuch vorstellen. Das DNS ist eine verteilte Datenbank, die einen Übersetzungsmechanismus zur Verfügung stellt. Ein für Menschen gut merkbarer Domänenname (z. B. "www.wikipedia.de") kann in eine IP-Adresse übersetzt werden kann und umgekehrt. Dieser Vorgang ereignet sich, unbemerkt für den Benutzer, immer dann, wenn er etwa im Webbrowser auf einen neuen Link klickt oder direkt eine Webadresse eingibt.

Die Standards und Protokolle des Internets werden in so genannten RFCs beschrieben und festgelegt.

Dienste

Im Internet finden sich unter anderem folgende Dienste:

- World Wide Web - per Hypertext verlinkte Webseiten und sonstige Medien, oft umgangssprachlich mit dem Internet gleichgesetzt
- Webforum - Diskussionen auf Webseiten
- WikiWikiWeb - offenes Autorensystem für Webseiten
- Weblog - einfaches Publizieren von Inhalten auf Webseiten
- E-Mail - der Postservice
- Mailingliste - Diskussionen per E-Mail
- Newsletter - Rundschreiben an Abonnenten
- File Transfer Protocol - Übertragung von Dateien
- Archie - Suchsystem für FTP-Server, weitgehend eingestellt
- Usenet - Diskussionsforen zu allen erdenklichen Themen
- Chat - Echtzeitkommunikation in Schriftform, z. B. im IRC oder als Instant Messaging
- Gopher - verteiltes Informationssystem, kaum noch in Gebrauch
- Veronica - Suchsystem für Gopher, ebenfalls kaum noch in Gebrauch
- WAIS - System zur Volltextsuche in verteilten Datenbeständen, auch kaum noch genutzt
- WAP - Technisch vereinfachte Version des World Wide Webs für Mobiltelefone
- i-mode - Weiterentwicklung von WAP mit zusätzlichen Funktionen
- Web Services - auf XML und HTTP basierende Dienste für Remote Procedure Calls
- Peer-to-Peer-Systeme - vor allem bekannt als Tauschbörsen zum Austausch von Dateien.Bekannte Vertreter sind z. B. eDonkey oder KaZaA.

Technik des Internet

Computernetzwerk

Ein **Computernetzwerk** ist ein Zusammenschluss von verschiedenen technischen Systemen (wie Computer, Handys, PDAs, Sensoren, Aktoren usw.) zum Zwecke der Kommunikation.

Unterschieden werden lokale Netzwerke (LAN, Intranet), nicht-lokale Netzwerke (MAN, WAN, GAN, Extranet), drahtgebundene Netzwerke (z. B. Ethernet) und drahtlose Netzwerke (Wireless LAN, Bluetooth, GSM, UMTS). Die Kommunikation erfolgt über verschiedene Protokolle, die mittels des ISO/OSI-Modells klassifiziert werden können.

Drahtgebundene Netzwerke

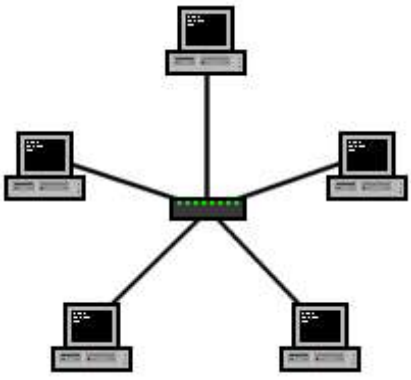
Verbreitete Techniken bei drahtgebundenen Netzwerken sind:

- Ethernet - größte Verbreitung
- Token Ring
- Token Bus
- FDDI - Glasfaserkabel

Topologien

Sterntopologie

Bei Netzen in **Sterntopologie** sind an einen zentralen Teilnehmer alle anderen Teilnehmer mit einer Zweipunktverbindung angeschlossen. Der zentrale Teilnehmer muß nicht notwendig über eine besondere Steuerungsintelligenz verfügen. In Transportnetzen ist das generell nicht der Fall. In Computernetzwerken kann es eine spezialisierte Einrichtung sein, zum Beispiel ein Sternkoppler, Hub oder Switch. Auch eine Telefonanlage ist gewöhnlich als Sternnetz aufgebaut: Die Vermittlungsanlage ist der zentrale Knoten an den die Teilnehmerapparate sternförmig angeschlossen sind.

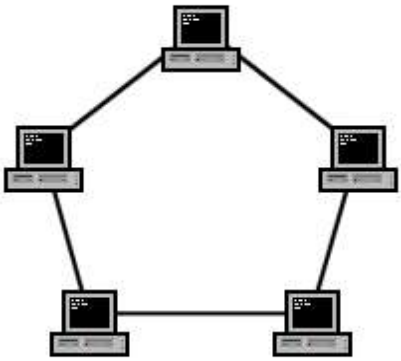


In jedem Fall bewirkt eine zentrale Komponente in einem Netz eine höhere Ausfallwahrscheinlichkeit für die einzelnen Verbindungen: ein Ausfall des zentralen Teilnehmers bewirkt unweigerlich den Ausfall aller Verbindungsmöglichkeiten zur gleichen Zeit.

Eine geläufige Schutzmaßnahme bei Sternnetzen besteht darin, die zentrale Komponente zu doppeln (Redundanz).

RINGTOPOLOGIE

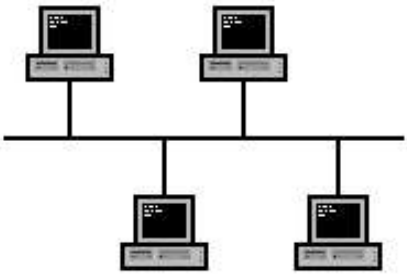
Bei der Vernetzung in **Ringtopologie** werden jeweils 2 Teilnehmer über Zweipunktverbindungen miteinander verbunden, so dass ein geschlossener Ring entsteht. Die zu übertragende Information wird von Teilnehmer zu Teilnehmer weitergeleitet, bis sie ihren Bestimmungsort erreicht. Um Überschneidungen zu verhindern, sind bei dieser Art der Vernetzung Adressierungsverfahren für die Information nötig. Da jeder Teilnehmer gleichzeitig als Repeater wirken kann (wenn keine Splitter eingesetzt werden) können auf diese Art große Entfernungen überbrückt werden (bei Verwendung von Lichtwellenleitern (LWL) im Kilometerbereich).



Bei einem Ausfall einer der Verbindung bricht das gesamte Netz zusammen, ausser die Teilnehmer beherschen Protection-Umschaltung. In einem Ring mit *Protection* wird häufig der Arbeitsweg in einer bestimmten Drehrichtung um den Ring geführt (z. B. im Uhrzeigersinn), der Ersatzweg in der anderen Drehrichtung (im Beispiel gegen den Uhrzeigersinn). Verwendung findet dieses Verfahren unter anderem auch bei Feldbussystemen auf Lichtwellenleiter-Basis.

BUSTOPOLOGIE

Eine **Bus-Topologie** besteht aus einem Hauptkabel, dem Bus, an das alle Geräte angeschlossen sind.



Der Anschluss zwischen den Geräten (also Netzwerkkarten) und Hauptkabel erfolgt über T-Verbinder. Unter Ethernet können bis zu vier Netzkarten im Server eingesetzt werden, wobei jede Karte über einen eigenen Transceiver zur Signalübertragung vom bzw. auf das Netzkabel verfügt.

Buszugriffsverfahren verhindern, dass sich die Teilnehmer gegenseitig stören. Sie regeln, welcher Teilnehmer die gemeinsame Leitung – den Bus – zu welchem Zeitpunkt zur Verfügung hat.

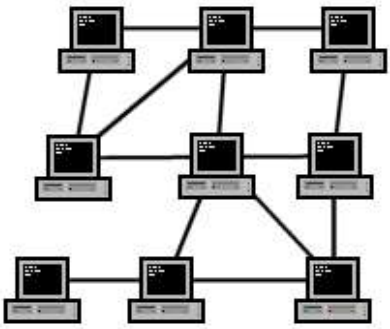
Bei diesem Verfahren treten folgende Probleme auf:

- Während des Datenverkehrs muss jeder Teilnehmer jede Sendung mithören. Dadurch steigt die Belastung (Strom) der Ausgangsbaugruppen des Senders mit der Anzahl der Teilnehmer am Bus.
- Feldbussysteme können sich über einen Bereich von mehreren hundert Metern erstrecken. Hier ist die Leitungslänge im Vergleich mit der Wellenlänge der Übertragung nicht mehr vernachlässigbar klein. Um störende Reflexionen zu vermeiden, werden Busabschlusswiderstände benötigt, die die Ausgänge des Senders ebenfalls mit höheren Strömen belasten. Kleinere Feldbussysteme können dennoch sehr gut nach dem Bus–Prinzip vernetzt werden.

Die Daten können in beide Richtungen übertragen werden. Vorteile eines Busnetzes sind der geringe Kabelbedarf und die Unabhängigkeit von der Funktion einzelner Stationen: Bei einem Ausfall eines Knoten oder einer Station bleibt das gesamte System trotzdem intakt. Größte Gefahr ist jedoch ein Kabelbruch im Hauptkabel, durch den der ganze Bus ausfällt.

Die Übertragungsrate bei beiden Leitungstypen (Thick Ethernet und Thin Ethernet) liegt bei 10-100 Mbit/s.

Vermaschtes Netzwerk



In einem **vermaschten Netzwerk** ist jedes Endgerät mit einem oder mehreren anderen Endgeräten verbunden. Wenn jeder Knoten mit jedem anderen Knoten verbunden ist, spricht man von einem vollständig vermaschten Netzwerk.

Bei Ausfall eines Endgerätes oder einer Leitung ist es im Regelfall möglich, durch umleiten (Routing) der Daten weiter zu kommunizieren.

In großen Netzwerken findet man oftmals eine Struktur, die sich aus mehreren verschiedenen Topologien zusammensetzt. So ist das Internet in weiten Teilen ein vermaschtes Netz, trotzdem gibt es "Hauptverkehrsadern" (die Backbone-Leitungen), die einem Bus ähneln.

Drahtlose Netze

Hier unterscheidet man Netze,

- die auf eine feste Infrastruktur angewiesen sind z. B.
- Mobilfunknetze wie GSM oder UMTS
- WLANs im Infrastruktur-Modus, d. h. mit Schnittstelle zu einem drahtgebundenen Netz mittels eines Access-Points. Am weitesten verbreitet sind WLANs vom Typ 802.11
- die spontan gebildet werden können, z. B. sogenannte MANETs
- 802.11-Netze im Ad-hoc-Modus. In diesem Modus kommunizieren die Geräte des Netzes ohne zusätzliche Infrastruktur.
- die mit sehr geringer Reichweite Geräte in unmittelbarer Umgebung verbinden, sogenannte WPANs

Sprachliche Betrachtung von Netz und Netzwerk

Das englische *net* wird traditionell in der Fischerei verwendet. Außerhalb dieses Bereichs wird daher *network* verwendet. Im Deutschen steht *Netzwerk* traditionell nur für das Maschenwerk eines Fischernetzes. Außerhalb der Fischerei wird nur *Netz* (Stromnetz, nicht

-werk; Telefonnetz) verwendet. Dieser Argumentation folgend ist *Computernetzwerk* eine falsche Übersetzung aus dem Englischen und *Computernetz* der korrekte Begriff.

Literatur

- Andrew S. Tanenbaum: *Computernetzwerke*. Pearson Studium 2003 ISBN 3827370469
- Douglas Comer: *Computernetzwerke und Internets*. Pearson Studium 2001 ISBN 382737023X

Netzwerkprotokoll

Ein **Netzwerkprotokoll** ist eine exakte Vereinbarung, nach der Daten über ein Computernetzwerk ausgetauscht werden. Eine solche Vereinbarung kann z. B. lauten: "Zunächst schickt Computer 1 die Zeichenfolge HELLO, anschließend schickt Computer 2 seine Adresse zurück, darauf sendet Computer 1 ein bestimmtes Kommando etc."

Mit Hilfe solcher Protokolle können Computer oder andere digitale Geräte zahlreiche Funktionen ausführen, z. B. Daten fehlerfrei zu einem anderen Computer befördern, E-Mails verschicken, Web-Seiten laden, etc. Diese Funktionen bauen zum Teil aufeinander auf. So löst beispielsweise das Protokoll TCP das Problem einer fehlerfreien Datenübertragung zu einem anderen Rechner. Das Protokoll SMTP zum Übermitteln von E-Mails benötigt selbst wiederum die Funktion, ein paar Zeichen zum anderen Rechner zu schicken und verwendet hierzu **TCP**. Diese Schichtung der Protokolle wird mit Hilfe des **OSI-Modells** dargestellt.

Beispiele für Netzwerkprotokolle sind die Internet-Protokolle (siehe auch TCP/IP-Referenzmodell) oder die Protokolle der AppleTalk-Familie.

In der Computertechnik gibt es noch eine Vielzahl anderer Protokolle die nicht immer Netzwerkprotokolle sind. Auch der Austausch von Daten zwischen CPU und RAM oder zwischen verschiedenen Peripheriegeräten wird durch Protokolle geregelt.

OSI-Modell

Das **OSI-Modell** (engl. *Open Systems Interconnection Reference Model*) ist ein offenes Schichtenmodell, das seit den 70er Jahren entwickelt und standardisiert wurde. Es teilt die verschiedenen Problembereiche der Netzwerkkommunikation in sieben Schichten auf, die aufeinander aufsetzen.

Weitere Bezeichnungen für das Modell sind *ISO/OSI-Modell*, *OSI-Referenzmodell*, *OSI-Schichtenmodell* oder *7-Schichten-Modell*.

Ein Netzwerk stellt seinen Benutzern Dienste bereit. Im einfachsten Fall überträgt es Daten von A nach B. Hierzu müssen jedoch tatsächlich eine Vielzahl von Aufgaben bewältigt werden. Die Probleme, die dabei gelöst werden müssen, reichen von Fragen der

elektronischen Übertragung der Signale über eine geregelte Reihenfolge in der Kommunikation (wer darf wann senden?) bis hin zu abstrakteren Aufgaben, die sich innerhalb der kommunizierenden Anwendungen ergeben. Die Vielzahl dieser Probleme und Aufgaben lässt es sinnvoll erscheinen, das Netz nicht als einen einzigen Dienstleister zu betrachten, sondern seine Dienste ganz bestimmten Kategorien zuzuordnen. Als besonders geeignet hat sich die Aufteilung in Schichten erwiesen.

Im OSI-Modell nimmt der Abstraktionsgrad der Funktionen von Schicht zu Schicht zu. Die Daten werden von einer Schicht zur nächsten weitergereicht, d. h. die Kommunikation erfolgt in vertikaler Richtung. Auf der Senderseite läuft die Kommunikation von oben nach unten und auf der Empfängerseite von unten nach oben.

Das Modell besteht aus sieben Schichten (engl. *layers*). Für jede Schicht sind die Dienste und Funktionen definiert, die auf ihr erfüllt werden sollen. Da jedoch keine Standards definiert sind, die diese Dienste und Funktionen verwirklichen, kann dies u.U. durch unterschiedliche Protokolle erfüllt werden.

Die 7 Schichten kann man sich gut mit dem Merksatz "**A**ll **p**eople **s**eem **t**o **n**eed **d**ata **p**rocessing" einprägen, wobei die Anfangsbuchstaben jeweils für die englische Bezeichnung stehen:

<table>
<tr><th colspan="2">OSI-Schicht</th><th>TCP/IP-Schicht</th><th>Kommunikation</th><th>Protokollbeispiele</th></tr>
<tr><td>7</td><td>Anwendung (Application)</td><td>Anwendung</td><td rowspan="4">Ende zu Ende (Multihop)</td><td rowspan="2">FTP
SMTP
HTTP</td></tr>
<tr><td>6</td><td>Darstellung (Presentation)</td><td rowspan="2">–</td></tr>
<tr><td>5</td><td>Sitzung (Session)</td><td rowspan="2">TCP</td></tr>
<tr><td>4</td><td>Transport</td><td>Transport</td></tr>
<tr><td>3</td><td>Netzwerk (Network)</td><td>Internet</td><td rowspan="3">Punkt zu Punkt</td><td>IP</td></tr>
<tr><td>2</td><td>Sicherung (Data Link)</td><td rowspan="2">Host an Netz</td><td rowspan="2">Ethernet, Token Ring, FDDI</td></tr>
<tr><td>1</td><td>Bitübertragung (Pysical)</td></tr>
</table>

Das OSI-Modell im Überblick (im Vergleich dazu das TCP/IP-Referenzmodell)

Die Protokollschichten

Anwendungsschicht, Schicht 7, die oberste Schicht

(engl. *application layer*, auch: Verarbeitungsschicht, Anwenderebene) Die Anwendungsschicht stellt den Anwendungen eine Vielzahl an Funktionalitäten zur Verfügung (z. B. Datenübertragung, E-Mail, Virtual Terminal bzw. Remote login etc.)

Darstellungsschicht, Schicht 6

(engl. *presentation layer*, auch: Datendarstellungsschicht, Datenbereitstellungsebene). Die Darstellungsschicht standardisiert die Datenstrukturen und ermöglicht somit den semantisch korrekten Datenaustausch zwischen unterschiedlichen Systemen (unter anderem Kodierung, Kompression, Kryptographie)

Sitzungsschicht, Schicht 5

(engl. *session layer*, auch: Kommunikationssteuerungsschicht, Steuerung logischer Verbindungen, Sitzungsebene) Um Zusammenbrüche der Sitzung und ähnliche Probleme zu beheben, stellt die Sitzungsschicht Dienste für einen organisierten und synchronisierten Datenaustausch zur Verfügung. Zu diesem Zweck werden so genannte Token eingeführt.

Transportschicht, Schicht 4

(engl. *transport layer*, auch: Ende-zu-Ende-Kontrolle, Transport-Kontrolle) Die Transportschicht ist die unterste Schicht, die eine vollständige Ende-zu-Ende Kommunikation (zwischen Sender und Empfänger) zur Verfügung stellt, d. h. für alle Schichten oberhalb der Netzwerkschicht ist die darunterliegende Netzwerktopologie transparent. Zu den Aufgaben der Transportschicht zählt die Segmentierung von Datenpaketen und die Stauvermeidung (engl. *congestion control*).

Netzwerkschicht, Schicht 3

(engl. *network layer*, auch: Vermittlungsschicht, Paketebene) Die Netzwerkschicht sorgt für die Weitervermittlung von Datenpaketen. Da nicht immer eine direkte Kommunikation zwischen Absender und Ziel möglich ist, müssen Pakete weitergeleitet werden. Weitervermittelte Pakete gelangen nicht in die höheren Schichten, sondern werden mit einem neuen Zwischenziel versehen und an den nächsten Host gesendet. Zu den Aufgaben der Netzwerkschicht zählt der Aufbau und die Aktualisierung von Routingtabellen, sowie die Flusskontrolle.

Sicherungsschicht, Schicht 2

(engl. *data link layer*, auch: Verbindungssicherungsschicht, Verbindungsebene, Prozedurebene) Aufgabe der Sicherungsschicht ist es, eine sichere (d. h. fehlerfreie) Verbindung zu gewährleisten und den Zugriff auf das Übertragungsmedium zu regeln. Daher teilt man die Schicht in zwei Subschichten auf: die LLC-Schicht (*logical link control*) und die Mediumzugriffsschicht (medium access control layer, MAC-Layer). Die Aufgaben der LLC-Schicht sind das Aufteilen des Bitdatenstromes in Datenrahmen (*frames*) und das Hinzufügen von Prüfsummen sowie das Verwalten von Quittungen und die Flusskontrolle. Die Mediumzugriffsschicht regelt konkurrierende Zugriffe mehrerer Stationen auf ein gemeinsames Übertragungsmedium und behandelt ggf. aufgetretene Kollisionen.

Physikalische Schicht, Schicht 1, die niedrigste Schicht

(engl. *physical layer*, auch: Bitübertragungsschicht, physikalische Ebene) Die physikalische Schicht ist für die eigentliche Bitübertragung der Daten zuständig. Hierzu ist eine Stan-

dardisierung der Netzwerk-Leitungen und -Anschlüsse sowie ihrer physikalischen Eigenschaften nötig. Die gemeinsame Nutzung eines Übertragungsmediums kann auf dieser Schicht durch ein statisches Multiplexing erfolgen.

Das OSI-Referenzmodell wird oft herangezogen, wenn es um das Design von Netzwerkprotokollen und die theoretische Betrachtung geht. Zusammen mit diesem Modell sind Netzwerkprotokolle entwickelt worden, die jedoch heute kaum eine Bedeutung besitzen. In der Praxis wird hauptsächlich die Familie der TCP/IP-Protokolle eingesetzt. Da das TCP/IP-Referenzmodell sehr speziell auf den Zusammenschluss von Netzen (*Internetworking*) zugeschnitten ist, bietet das OSI-Referenzmodell einen umfassenderen Ansatz für die Betrachtung von Netzwerkprotokollen.

Das OSI-Modell betreffende Standards:

- ISO 7498-1 (DIN ISO 7498)
- ITU-T (CCIT) X.200

Literatur

- Stahlknecht, P./Hasenkamp, U.; *Einführung in die Wirtschaftsinformatik*; Springer; Berlin; 2002

Internet-Protokoll-Familie

Die **Internet-Protokoll-Familie** (engl: *internet protocol suite*) ist eine Familie von Netzwerkprotokollen, die die Basis für die Netzwerkkommunikation im Internet bilden. Synonym dazu werden auch die Bezeichnungen DoD- oder TCP/IP-Protokoll-Familie verwendet. *DoD* (engl. *Department of Defense* "US-Verteidigungsministerium") hat die Entwicklung der Internet-Protokolle im Rahmen des *DARPA Internetwork Project* finanziert. TCP/IP steht für zwei Protokolle, die das Herzstück dieser Protokoll-Familie bilden: Das Transmission Control Protocol und das Internet Protocol.

Entstehung

Am 1. Januar 1983 wurde im damaligen Arpanet, dem Vorläufer des Internet, das veraltete NCP (Network Core Protocol) gegen TCP/IP ausgetauscht, und somit die Basis für das heutige Internet geschaffen. Das Protokoll wurde von Vinton Cerf und Robert E. Kahn entwickelt und setzte sich anfangs nur zögerlich gegen andere Lösungen wie das offizielle OSI (Open System Interconnection Modell) der ISO durch. Der entscheidende Umschwung für TCP/IP kam dann Mitte der neunziger Jahre, als mit der Seitenbeschreibungssprache HTML und dem Browser Mosaic für breite Anwendungen taugliche Internetwerkzeuge zur Verfügung standen.

<table>
<tr><td>Anwendung</td><td>FTP</td><td>SMTP</td><td>HTTP</td><td>DNS</td><td>...</td></tr>
<tr><td>Transport</td><td colspan="3">TCP</td><td colspan="2">UDP</td></tr>
<tr><td>Netzwerk</td><td colspan="4">IP (IPv4,IPv6)</td><td>ARP</td></tr>
<tr><td rowspan="2">Netzzugang</td><td rowspan="2">Ethernet</td><td rowspan="2">FDDI</td><td rowspan="2">Token Ring</td><td>PPP</td><td rowspan="2">...</td></tr>
<tr><td>Modem</td></tr>
</table>

Internet Protokolle (Protokollstapel)

Die derzeit aktuelle IP Version 4 stößt mittlerweile an ihre Grenzen. Die Anzahl der Netzadressen ist zu stark begrenzt und die Echtzeitübertragung von Video und Sprache funktioniert nur unzureichend, da TCP/IP Version 4 keine Verfahren zur Festlegung der Bandbreite, Geschwindigkeit und Priorität hat, was als Quality of Service (QoS) bezeichnet wird. Deswegen wurde das Internet Protocol weiterentwickelt und wird als IPv6 in den nächsten Jahren eingeführt. Ein weiteres Merkmal, das den Erfolg der IP-Protocol-Suite ermöglichte, ist die offene Definiton der Zugriffsschicht. Diese entspricht den OSI-Schichten 1 (Physikalische Schicht) und 2 (Sicherungsschicht). Daher steht TCP/IP für beinahe jede Schicht 1/2 Technologie zur Verfügung. Beispiele: Gigabit-/Fast-/Ethernet, Token Ring, FDDI, PPP, X.25, Frame Relay, aber auch ATM, Sonet, SDH.

Internet-Protokollfamilie im Detail

Zur Gliederung der Kommunikationsaufgaben werden in Netzwerken funktionale Ebenen, so genannte Schichten, unterschieden. Für die Internetprotokolle ist das TCP/IP-Referenzmodell maßgebend. Die Internetprotokolle beginnen dabei bei Schicht3 mit dem für das Weiterleiten von Paketen in verschiedene Netze zuständigen Internet Protocol (IP, meist IPv4). Internet Protokolle werden in den RFC-Dokumenten diskutiert und standardisiert. Die wesentlichen Protokolle haben dabei den Status erforderlich (*required*), empfohlen (*recommended*), oder optional (*elective*). Andere Protokolle können auch als experimentell (*experimental*) oder historisch (*historic*) bezeichnet werden.

Erforderliche Protokolle

IPv4 und ICMP müssen in jeder TCP/IP-Implementation vorhanden sein. Ein Router kann nur diese beiden Protokolle verwenden. Aus Anwendungssicht ist damit aber noch nichts Sinnvolles möglich. Dazu werden noch empfohlene und weitere Protokolle benötigt.

Die verbreitete IP-Version 4 (IPv4), früher einfach als *IP-Protokoll* bezeichnet, wurde schon 1981 definiert und ist im Wesentlichen unverändert noch heute auf allen System im Internet das hauptverwendete Internet-Protokoll. Die neue Version 6 IPv6 soll die Version 4 nach und nach ablösen. Das ICMP-Protokoll wird für die Fehlersuche und für Informationsmeldungen verwendet.

Empfohlene Protokolle

Vor jedes Anwendungsprotokoll (oberste Schicht) wird stets eines der beiden Transportprotokolle TCP (*Transport Control Protocol*) für verbindungsorientierte und UDP (*User Datagram Protocol*) für paketorientierte Anwendungen eingeschoben (Schicht 4)

Für eine TCP/IP-Implementation werden folgende Anwendungsprotokolle ebenfalls empfohlen:

- TELNET (Teletype over the Network) - Einloggen in ein Terminal über das Netzwerk (TCP-Transport)
- FTP (File Transfer Protocol) - Dateiübertragung (TCP-Transport)
- SMTP (Simple Mail Transfer Protocol) - Elektronische Post E-Mail (TCP-Transport)
- DNS (Domain Name Service) - Namen- zu Adressübersetzung, E-Mail Routing (TCP und UDP)
- ECHO (Echo Protocol) - Fehlersuche, Antwortzeiten (TCP und UDP)
- NTP (Network Time Protocol) - Zeitsynchronisation (UDP)
- SNMP (Simple Network Management Protocol) - Netwerküberwachung und -management (UDP)
- BOOTP (Boot Protocol) - Parameter für plattenlose Geräte (UDP)

Optionale Protokolle

Weitere oft verwendete Protokolle sind HTTP für das WWW, DHCP für die Rechnerkonfiguration, POP3 für den Mailempfang und SSH für sicheres Einloggen (Ersatz für TELNET). ARP wird zur dynamischen Auflösung von Hardwareadressen im Ethernet benötigt.

Welche Protokolle auf welchen Rechnern

- (Kleine) Router implementieren z. B. ARP, IP, ICMP, UDP, SNMP, RIP.
- WWW-Clients verwenden ARP, IP, ICMP, UDP, TCP, DNS, HTTP, FTP.
- Benutzerrechner verwenden zusätzlich z. B. TELNET, SMTP, POP3, SNMP, ECHO, DHCP, SSH, NTP.

Plattenlose Geräte implementieren im ROM ARP, IP, ICMP, UDP, BOOTP , TFTP (nur paketorientierte Protokolle)

Literatur

- Douglas E. Comer: *Internetworking with TCP/IP. Principles, Protocols, and Architectures.* Prentice Hall, 1995, ISBN 0-13-018380-6
- Richard Stevens: *TCP/IP Illustrated I. The Protocols.* Addison-Wesley Professional 1994 ISBN 0201633469

Internet Protocol

Das Internet Protocol bildet die erste vom Übertragungsmedium unabhängige Schicht der Internet-Protokoll-Familie. Im Gegensatz zu der physikalischen Adressierung der darunter liegenden Schicht, bietet IP *logische* Adressierung. Das bedeutet, dass mittels IP-Adresse und Subnetzmaske (*subnet mask*) Computer innerhalb eines Netzwerkes in logische Einheiten, so genannte Subnetze, gruppiert werden können.

Anwendung	FTP	SMTP	HTTP	DNS	...
Transport	TCP			UDP	
Netzwerk	**IP (IPv4, IPv6)**				
Netzzugang	Ethernet	Token Bus	Token Ring	FDDI	...

Internet Protokolle im TCP/IP-Protokollstapel

Auf dieser Basis ist es möglich, Computer in größeren Netzwerken zu adressieren und Verbindungen zu ihnen aufzubauen, da logische Addressierung die Grundlage für Routing (Wegewahl und Weiterleitung von Netzwerk-Paketen) ist. IP stellt also die Grundlage des Internets dar.

IPv4

IPv4, früher einfach **IP**, ist die vierte Version des Internet Protocols (IP). Es war die erste Version des Internet Protocols das weit verbreitet und eingesetzt wurde und bildet eine wichtige technische Grundlage des Internets. Es wurde in RFC 791 im Jahr 1981 von Jon Postel definiert.

Adressformat

IPv4 benutzt 32-Bit Adressen, d. h. maximal sind 4.294.967.296 eindeutige Adressen möglich. IPv4-Adressen werden dezimal geschrieben (z. B. 80.130.234.185). Bei der Weiterentwicklung IPv6 werden 128-Bit Adressen verwendet. Es ist aber heute noch nicht sehr verbreitet.

Eine IP-Adresse wird in einen Netzwerkteil und einen Host-(Adressen-)teil getrennt. Ein typisches Netzwerk trennt die 32-Bit in einen 24-Bit Netzwerk und einen 8-Bit Adressteil (Klasse-C Adressen). Der Adressteil wird im LAN (lokales Netzwerk) vergeben, der Netzwerkteil von der IANA zugeteilt. In einem Klasse-C Netzwerk sind 254 Geräteadressen verfügbar. Die Adresse 255 wird für Broadcast, d. h. Übertragung an alle, verwendet, die Adresse 0 für das Netzwerk selbst. Eine Netzmaske, im Klasse-C Fall 255.255.255.0 gibt die Trennlinie zwischen den Teilen an. Sie ist 1 für den Netzwerkteil und 0 für den Adressteil.

Ein Teil der Adressen ist für den lokalen Rechner (127.x.x.x), für Multicasts (224.x.x.x) und für lokale Netzwerke mit Adressübersetzung (z. B. 10.x.x.x und 192.168.x.x), sowie allgemeiner Broadcast (255.255.255.255) reserviert.

Manchmal teilt der Eigentümer den ihm zugeteilten Adressteil lokal in weitere Subnetze auf (so genanntes *Subnetting*). Dies dient der besseren Netzwerkausnutzung sowie der Fehlersuche.

Paketlänge

Ein IP-Paket besteht aus einem Header und den eigentlichen Daten. Der Datenteil enthält in der Regel ein weiteres Protokoll, meist TCP, UDP oder ICMP. Die maximale Länge der Daten beträgt 65515 Bytes (2^{16}-1-*minimale Headerlänge*). Normalerweise beschränkt der Sender die Paketlänge auf diejenige des zugrundeliegenden Mediums. Bei Ethernet beträgt die so genannte MTU (*Maximum Transfer Unit*) 1518 Bytes, wobei 18 Bytes von Ethernet selbst belegt werden. Für IP (Header und Daten) stehen also nur 1500 Bytes zur Verfügung.

Andere Netzwerke können die Paketlänge weiter beschränken. In diesem Fall bietet IP die Option, IP-Pakete zu fragmentieren. Jedes Paket erhält vom Sender eine Kennung (die Fragment-ID). Ein Router kann ein langes Paket aufteilen, weil das Zielnetzwerk Pakete dieser Länge nicht überträgt. Der Empfänger kann die Fragmente anhand der Kennung und der Senderadresse identifizieren und wieder zusammenfügen.

Routing

IPv4 unterscheidet nicht zwischen Endgeräten (Hosts) und Vermittlungsgeräten (Router). Jeder Computer und jedes Gerät kann gleichzeitig Endpunkt und Router sein. Ein Router verbindet dabei verschiedene Netzwerke. Die Gesamtheit aller über Router verbundenen Netzwerke bildet das Internet.

IPv4 ist für LANs und WANs gleichermaßen geeignet. Ein Paket kann verschiedene Netzwerke vom Sender zum Empfänger durchlaufen, die Netzwerke sind durch Router verbunden. Anhand von Routingtabellen, die jeder Router individuell pflegt, wird der Netzwerkteil einem Zielnetzwerk zugeordnet. Die Einträge in die Routingtabelle könnrn dabei statisch oder über Routingprotokolle dynamisch erfolgen. Die Routingprotokolle dürfen dabei sogar auf IP aufsetzen.

Bei Überlastung eines Netzwerks oder einem anderen Fehler darf ein Router Pakete auch verwerfen. Pakete desselben Senders können bei Ausfall eines Netzwerks auch alternativ geroutet werden. Jedes Paket wird dabei einzeln geroutet, was zu einer erhöhten Ausfallsicherheit führt.

Beim Routing über IP können daher

- einzelne Pakete verlorengehen
- Pakete doppelt beim Empfänger ankommen
- Pakete verschiedene Wege nehmen
- Pakete fragmentiert beim Empfänger ankommen.

Wird TCP auf IP aufgesetzt (d. h. die Daten jedes IP-Pakets enthalten ein TCP-Paket, aufgeteilt in TCP-Header und Daten), so wird neben dem Aufheben der Längenbeschränkung auch der Paketverlust durch Wiederholung korrigiert. Doppelte Pakete werden erkannt und verworfen. Die Kombination TCP mit IP stellt dabei eine zuverlässige bidirektionale Verbindung eines Datenstroms dar.

ICMP

IP ist eng verknüpft mit dem ICMP-Protokoll, das zur Fehlersuche und Steuerung eingesetzt wird. ICMP setzt auf IP auf, d. h. ein ICMP-Paket wird im Datenteil eines IP-Pakets abgelegt. Eine IP-Implementierung enthält stets auch eine ICMP Implementierung. Wichtig ist z. B. die ICMP Source-Quench Mitteilung, die den Sender über das Verwerfen von Paketen durch einen Router informiert. Da jedes IP-Paket die Quell-Adresse enthält, können Informationen an den Sender zurück übermittelt werden. Dieser kann nach einem Source-Quench die Paketsendefrequenz verringern und so die Notwendigkeit eines weiteren Verwerfens minimieren oder vermeiden.

ICMP kann zusammen mit dem *Don't Fragment*-Bit des IP-Pakets auch eingesetzt werden, um die minimale MTU eines Übertragungsweges zu ermitteln (so genannte PMTU *Path Maximum Transfer Unit*). Dies ist die MTU desjenigen Netzwerkes mit der kleinsten MTU aller passierten Netzwerke. Dadurch kann auf Fragmentierung verzichtet werden, wenn der Sender nur Pakete mit der maximalen Größe der PMTU erzeugt.

IPv4 auf Ethernet

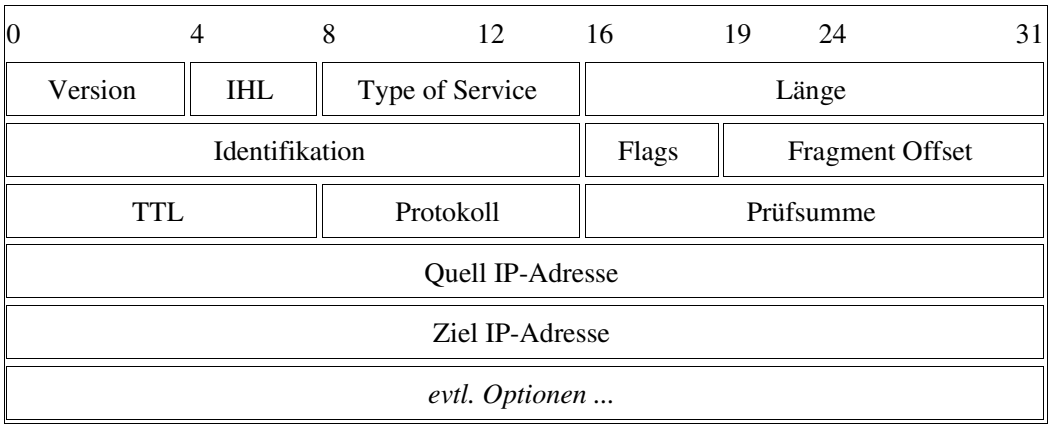
IPv4 kann auf vielen verschiedenen Medien aufsetzen, z. B. auf serielle Schnittstellen (PPP oder SLIP-Protokoll), Satellitenverbindungen usw. Im LAN-Bereich wird heute fast immer Ethernet eingesetzt. Ethernet verwaltet eigene 48-Bit Adressen. Wenn IP über Ethernet gesendet wird, wird ein 14-Byte grosser Ethernet-Header vor dem IP-Header gesendet. Nach den Daten folgte eine 32-Bit CRC-Prüfsumme. Neben der maximalen Paketlänge von 1518 Bytes kann Ethernet keine kleineren Pakete als 64 Bytes übertragen, so dass zu kurze IP-Pakete (Datenlänge kleiner als 26 Bytes) mit Nullbytes erweitert werden (so genanntes *Padding*). Die Länge im IP-Header gibt dann Auskunft über die tatsächliche Paketgröße.

Im Ethernet hat jede Netzwerkkarte hat ihre eigene herstellerbezogene 48-Bit Adresse, zusätzlich gibt es eine eigene Broadcastadresse. Ein Sender muss die Ethernetadresse der Zielnetzwerkkarte kennen, bevor ein IP-Paket gesendet werden kann. Dazu wird das ARP-Protokoll (*Address Resolution Protocol*) verwendet. Jeder Rechner verwaltet einen ARP-Cache, in dem er ihm bekannte Zuordnungen von Ethernet-Kartenaddressen speichert. Unbekannte Adressen erfährt er über das ARP-Protokoll mittels einer Anfrage (ARP-Request) über einen Ethernet-Broadcast, die der zugehörige Empfänger beantwortet (ARP-Reply).

Header Format

Der IPv4 Header ist normalerweise 20 Bytes lang. Bei Übertragung auf Ethernetkabeln folgt er dem Ethernet-Typfeld, das für IP-Pakete auf 0800_{16} festgelegt ist. Auf anderen Übertragungsmedien kann der Header auch der erste Eintrag sein.

IPv4 bietet verschiedene, größtenteils ungenutzte Optionen, die den Header bis auf 60 Bytes (in 4-Byte Schritten) verlängern können.



Da das *Type of Service*-Feld nie richtig genutzt wurde, wird es in modernen Implementationen für DiffServ und Explicit Congestion Notification verwendet. Einige nicht standardkonforme Firewalls verlassen sich jedoch darauf, dass dieses Feld gleich Null ist und sehen ein Paket andernfalls als "böse" an. Dies führt dazu, dass manche Subnetze für ECN/DiffServ-fähige Clients nicht erreichbar sind.

Höhere Protokolle

IPv4 ist ein Routingprotokoll (Schicht 3 im TCP/IP-Referenzmodell). Auf IPv4 werden weitere Protokolle aufgesetzt, d. h. in den Datenteil des IP-Pakets werden die Header, Daten und evtl. Trailer der oberen Protokolle eingefügt (Protokollstapel).

Neben dem erwähnten ICMP wird TCP verwendet, das TCP/IP zusammen mit IP den Namen gegeben hat. TCP ist ein verbindungsorientiertes Protokoll, das einen byteorientierten, bidirektionalen, zuverlässigen Datenstrom zur Verfügung stellt. Es wird im WAN-Bereich praktisch immer alle Arten von Daten- und Informationsübertragungen eingesetzt.

Zu TCP/IP gehört auch UDP, ein paketorientiertes Protokoll. Es ist ein einfaches Protokoll, das die Paketeigenschaften von IP im wesentlichen beibehält (verbindungslos, unzuverlässig, Verdoppelung etc.). TCP und UDP fügen IP eine Prüfsumme über die Daten (die Prüfsumme im IP-Header prüft nur die Headerdaten), und als Quell- und Zielport jeweils eine 16-Bit Zahl hinzu. Diese Ports bilden zusammen mit der jeweiligen Quell- und Zieladresse im IP-Paket so genannte Endpunkte. Prozesse kommunizieren über diese Endpunkte. TCP baut eine Verbindung nicht zwischen IP-Adressen, sondern zwischen zwei Endpunkten auf. Die weiteren Protokolle setzen alle entweder auf TCP oder auf UDP auf. Ein wichtiges Protokoll ist das *Domain Name System* DNS, das eine Umsetzung von Rechnernamen zu IP-Adressen erlaubt. Es überträgt Informationen normalerweise über UDP, der Abgleich zwischen zwei DNS-Servers kann aber auch das TCP-Protokoll verwenden.

Vergangenheit und Zukunft von IPv4

Das IPv4-Protokoll hat lange nahezu unverändert überlebt. Ab 1983 wurde die IP-Protokoll-Familie als einzige Protokollfamilie für das ARPAnet übernommen, das dann später zum Internet wurde. Damals waren nur einige hundert Rechner an das Netz angeschlossen. 1989 wurde die Grenze von 100.000 Rechnern überschritten, und im gleichen Jahr der Backbone auf 1.5 MBit/s aufgerüstet. Anfang der 1990er Jahre wurde vermutet, dass die IP-Adressen am Ende des Jahrzehnts knapp würden, da die IANA relativ große Adressbereiche an Firmen und Institutionen zuteilte. Dies führte zuerst zur Entwicklung eines Entwurfes für einen Standard (IPv5), der dann aber zu Gunsten von IPv6 verworfen wurde. IPv5 sollte dabei einen 64Bit Adressbereich unterstützen. Die gestiegenen Sicherheitsanforderungen und die Verwendung von Firewalls mit *Network Addess Translation.* NAT führte jedoch zu einer Entschärfung der Adressproblematik, da heute ganze Firmen mit mehreren tausend Rechnern nur noch eine einzige IP-Adresse pro Standort benötigen. Einige Eigenschaften, wie Fragmentierung, werden nicht mehr benötigt, da sie für die heutigen schnellen Netze zu aufwändig sind. *Path Maximum Transfer Unit Discovery* löst dieses Problem. IPv4 wird auch in nächster Zukunft noch das allgemein verwendete Protokoll im Internet bleiben. Schließlich hat IP auch die konkurrierenden LAN-Protokolle wie DECnet verdrängt. Netware, AppleTalk und NETBIOS wurden als neue Versionen hervorgebracht, die auf IP aufsetzen

IPv6

IPv6, das **Internet Protocol Version 6**, ist der Nachfolger des gegenwärtig im Internet noch fast ausschließlich verwendeten Internet Protocol v4.

Warum ein neues Internet-Protokoll?

Das alte IPv4 bietet einen Adressraum von etwas über 4 Milliarden IP-Adressen, mit denen Computer angesprochen werden können. In den Anfangstagen des Internet, als es nur wenige Rechner gab, die eine IP-Adresse brauchten, galt dies als mehr als ausreichend. Kaum jemand konnte sich vorstellen, dass überhaupt jemals so viele Rechner zu einem einzigen Netzwerk zusammengeschlossen würden, dass es im vorgegebenen Adressraum eng werden könnte.

Viele der theoretisch 4 Milliarden IP-Adressen sind in der Praxis nicht nutzbar, da sie Sonderaufgaben dienen (z. B. Multicast) oder zu großen Subnetzen gehören: Den ersten großen Teilnehmern am Internet wurden riesige Adressbereiche (sogenannte Class-A-Netze) mit je 16,7 Millionen Adressen zugeteilt, die diese Organisationen bis heute behalten haben, ohne sie jemals voll ausnutzen zu können. Die Amerikaner (und teilweise die Europäer) teilten die relativ wenigen großen Adressbereiche unter sich auf, während die Internet-Späteinsteiger wie Südamerika, aber vor allem Asien, zunächst außen vor blieben. Als Resultat herrscht besonders im zukünftigen IT-Wachstumsmarkt Asien heute eine latente Adressenknappheit, der man mit Notbehelfen wie NAT (Network Adress Translation) oder dynamischer Vergabe von Adressen begegnen muss.

Auf Grund des Wachstums und der Wichtigkeit des Internet konnte dies kein Dauerzustand bleiben; Zusätzlich ist abzusehen, dass in den nächsten Jahren durch neue technische Innovationen (z. B. Handys mit Internet-Anschluss, bald wohl auch Fernseher, Mikrowellen, Kühlschränke und Autos) der Bedarf an Adressen auch im Rest der Welt ansteigen wird.

Hauptsächlich wegen der Adressknappheit, aber auch, um einige der Probleme zu lösen, die sich im Zuge der großräumigen Verwendung von IPv4 gezeigt hatten, begann man 1995 mit den Arbeiten am neuen IPv6 (die ersten RFCs waren 1883ff.). Folgende Liste soll einen kurzen Überblick über die wesentlichen neuen Features von IPv6 geben, einige Punkte werden weiter unten näher erklärt:

- Vergrößerung des Adressraums von 2^{32} bei IPv4 auf 2^{128} bei IPv6
- Autokonfiguration (ähnlich DHCP), Mobile IP und automatisches Renumbering
- Services wie IPsec, QoS und Multicast serienmäßig
- Vereinfachung und Verbesserung der Header (wichtig für Router)

Adressaufbau von IPv6

Eine IPv6-Adresse ist 128 bit lang (IPv4: 32 bit). Damit gibt es etwa $3{,}4 \times 10^{38}$ IPv6-Adressen - für jeden Quadratmeter Erdoberfläche könnten $6{,}5 \times 10^{23}$ Adressen bereitgestellt werden. IPv6-Adressen werden nicht mehr dezimal (z. B. 80.130.234.185), sondern hexadezimal mit Doppelpunkten geschrieben: *243f:6a88:85a3:08d3:1319:8a2e:0370:7344*. Wenn eine 16-bit-Gruppe den Wert 0000 hat, kann sie durch einen Doppelpunkt ersetzt werden. Wenn dann mehr als 2 Doppelpunkte aufeinander folgen würden, können diese auf 2 Doppelpunkte reduziert werden, solange es in der resultierenden Adresse nur einmal zwei

aufeinander folgende Doppelpunkte gibt. *0588:2353::1428:57ab* ist also das selbe wie *0588:2353:0000:0000:0000:0000:1428:57ab*, aber *3906::25de::cade* wäre nicht erlaubt, da zweimal zwei Doppelpunkte in der Zeichenkette vorkommen - ein Computer wüsste nicht, wo mit wie vielen Nullen aufzufüllen wäre.

Die ersten 64 Bit der IPv6-Adresse dienen standardmäßig der Netzadressierung, die letzten 64 Bit können zur Host-Adressierung verwendet werden - hiermit implementiert man elegant das Konzept der Netzmasken von IPv4. Wer möchte, kann jedoch auch andere Netzmasken verwenden.

Arten von IPv6-Adressen

Es gibt verschiedene IPv6-Adressen mit Sonderaufgaben und unterschiedlichen Eigenschaften. Diese werden durch die ersten Bits der Adresse (das "Präfix") signalisiert:

- Das Präfix *00* steht für IPv4 und IPv4-über-IPv6-Kompatibilitätsadressen. Ein geeigneter Router kann diese Pakete zwischen IPv4 und IPv6 konvertieren und so die neue mit der alten Welt verbinden. Zwei weitere Adressen tragen ebenfalls dieses Präfix; *::0* ist die undefinierte Adresse, ähnlich der 0.0.0.0 in IPv4, und *::1* ist die Adresse des Loopback Devices "localhost".
- Die Präfixe 2 oder 3 stehen für Globale Unicast-Adressen, also eine routbare und weltweit einzigartige Adresse.
    - Eine Abart davon sind die 6Bone-Testadressen *3ffe*, die einem Rechner gehören, der Teil des IPv6-Testnetzwerkes 6Bone ist.
    - *2001*-Adressen werden an Provider vergeben, die diese dann wieder an ihre Kunden verteilen
    - *2002*-Präfixe deuten auf Adressen des Tunnelmechanismus 6to4 hin.
- *fe80* bis *febf* sind so genannte Link-local-Adressen, die nicht geroutet werden dürfen und daher nur im gleichen Subnetz erreichbar sind. Interessant werden sie bei der Autokonfiguration.
- Adressen mit Präfixen *fec0* bis *feff* sind die Nachfolger der privaten IP-Adressen (z. B. 192.168.x.x). Sie dürfen nur innerhalb der gleichen Organisation geroutet werden. Man nennt sie daher Site-local.
- Das Präfix *ff* steht für Multicast-Adressen. Dem Präfix folgt eine Angabe über den Gültigkeitsbereich des Pakets:
    - *ffx1*: Node-lokal, diese Pakete verlassen den Knoten nie.
    - *ffx2*: Link-lokal, werden von Routern grundsätzlich nie weitergeleitet und können deshalb das Subnetz nicht verlassen.
    - *ffx5*: Site-lokal, dürfen zwar geroutet werden, jedoch nicht von Border-Routern.
    - *ffx8*: Organisations-lokal, die Pakete dürfen auch von Border-Routern weitergeleitet werden, bleiben jedoch "in der Firma" (hierzu müssen seitens des Routing-Protokolls entsprechende Vorkehrungen getroffen werden).

- *ffxe*: Globaler Multicast, der nach überall hin geroutet werden darf.

Autokonfiguration

Ein IPv6-fähiges Interface kann aus seiner Layer 2-MAC-Adresse eine sogenannte Link-lokale Adresse errechnen, mit der es sich auf die Suche nach den Routern in seinem Netzwerksegment machen kann. Der Router kann dem Gerät dann eine "Unicast"-Adresse aus seinem Adressbereich zuweisen, mit der das Gerät aufs Internet zugreifen kann. Der ganze Vorgang läuft ohne Benutzerintervention vollautomatisch ab und ist eine Verbesserung des IPv4-DHCP (er kommt ohne Server aus). Ein IPv6-fähiges Gerät ist so "out-of-the-box" startklar, was besonders für unerfahrene Endnutzer oder stressgeplagte Admins ein großer Vorteil ist.

Renumbering

Sobald man bei IPv4 erstmal genug Rechner zu einem Subnetz zusammengeschlossen hat, bekommt man Probleme, wenn sich an den Adressparametern dieses Netzes etwas ändert (z. B. ein Providerwechsel). Jeder Rechner darf dann nämlich per Hand umkonfiguriert werden - wenn das Netz groß genug ist, kann sich dies als praktisch undurchführbar darstellen. In diesem Fall hat man bei IPv4 häufig den Adressraum fragmentiert, d. h. ein Teil der IP-Adressen eines Subnetzes wurde anders geroutet als der Rest. Für die Router brachte dies eine vergrößerte Routing-Tabelle, da sie sich die Ausnahmen merken und beachten müssen, und damit Performance-Verlust. Bei IPv6 gibt es diese Probleme nicht, da man ja den Autokonfigurationsmechanismus hat. Der neue Adressbereich muss nur einmal neu am Router eingestellt werden, und schon haben alle Clients ihre neuen Adressen.

Mobile IP

Bei Mobile IP geht es darum, mit der gleichen IP-Adresse überall erreichbar zu sein, beispielsweise im heimischen Netzwerk und auf einer Konferenz. Normalerweise müssten dazu Routing-Tabellen geändert werden (noch dazu für jede mobile IP-Adresse einzeln!). IPv6 regelt dies unter Zuhilfenahme eines Agenten-Rechners am Hauptstandort des Mobilgeräts und ICMPv6-Redirects. Eingehende Verbindungen werden durch den Agenten an den momentanen Standort des Mobilgeräts umgeleitet.

Effizienzsteigerungen und neues Header-Format

Im Gegensatz zu IPv4 hat der IP-Header bei v6 eine feste Länge. Außerdem gibt es keine Paketfragmentierung mehr und es werden keine Checksummen über das IP-Paket berechnet, man nutzt nur noch die Fehlerkorrektur in den Layern 2 und 4. Die meisten Felder im Header sind auf 64-Bit-Grenzen ausgerichtet, so dass der Speicherzugriff im Router nicht unnötig ausgebremst wird. Die Flags liegen jetzt in Zwischenheadern zwischen dem IP-Header und dem Layer-4-Header (TCP/UDP). Dadurch können sich Hochleistungsrouter auf ihre

eigentlichen Kernaufgaben konzentrieren und müssen nicht mehr die Fehler anderer Schichten ausbügeln.

Der frühere Broadcast wird nun konsequent über Multicast implementiert: ARP wurde beispielsweise durch das neue Verfahren "Neighbor Solicitation" ersetzt: Jeder Host muss sich bei einer aus seiner IP gebildeten Multicastgruppe anmelden, an die der anfragende Host seine Neighbor-Solicitation-Anfrage schickt und so die MAC-Adresse des anderen Hosts herausfinden kann. Als Nebeneffekt werden diese Anfragen nicht mehr im ganzen LAN verteilt, sondern gehen in geswitchten Umgebungen idealerweise nur noch über den direkten Weg zum eigentlichen Ziel.

Pferdefüsse

Gerade zu Anfang litt IPv6 unter einigen Kinderkrankheiten, die dem neuen Protokoll den Ruf einer "Totgeburt" einbrachten. Die Standards wurden häufig geändert, was einigen Unmut erzeugte, da man die gerade fertig gewordenen Implementationen schon wieder wegwerfen konnte. Der größte Einschnitt bestand in der Einführung der IEEE-Norm EUI-64 für die Link-local-Adressen. Vorher übernahm man die MAC-Adresse des Adapters einfach in die IPv6-Adresse, nun wurde die MAC-Adresse gemäß EUI-64 in veränderter Form in die IPv6-Adresse übernommen. Das änderte jedoch nichts an dem Problem, dass aus der gleichen MAC-Adresse auch immer die gleiche IPv6-Adresse resultiert. Dynamische IP-Vergabe wie bei IPv4 sollte es ja bei IPv6 nicht mehr geben. Datenschützer waren besorgt, dass auf diese Weise der Datenverkehr einer IP auf Routern mitgeschnitten werden könnte und z. B. für Marketing-Maßnahmen oder staatliche Interventionen aller Art verwendet werden könnte. Die IETF definierte deshalb nachträglich die Privacy Extensions gemäß RFC 3041: Die MAC-Adresse wird dabei zunächst mit einer pseudo-zufälligen Zahl verwürfelt, und aus dem Ergebnis dann die Link-lokale Adresse des Gerätes ermittelt.

In der Praxis ergab sich mit den Link-local-Adressen das Problem, dass es nicht mehr reicht, die IP-Adresse im Destination-Feld einzutragen, sondern auch eine Scope-ID angegeben werden muss, da die Link-local-Adressen relativ zum Link sind und alleine noch keinen Endpunkt definieren.

Deshalb sind die Link-local-Adressen nur beschränkt zur Kommunikation tauglich (abhängig davon, ob die IPv6-Unterstützung der verwendeten Anwendung das Konzept der Scope-ID kennt).

Die Praxis: IPv6-Verfügbarkeit

IPv6 setzt sich im praktischen Einsatz nur langsam durch. Mit der Betriebssystemunterstützung sieht es im Moment folgendermaßen aus:

- **Windows Server 2003**: Enthält auch einen "Production Quality" Stack und unterstützt IPv6 DNS (AAAA) Einträge und IPv6 Routing.

- **Windows XP**: Auf expliziten Wunsch (*ipv6 install*) kann man bei Windows XP einen experimentellen IPv6-Stack installieren. Ab ServicePack 1 hat dieser Stack "Production Quality", muss aber immer noch explizit installiert werden.
- **Windows 2000**: Microsoft bietet einen experimentellen Stack als Patch an.
- **Windows 9x/ME**: Lediglich eine kommerziell verfügbare Unterstützung der Firma Trumpet (Winsock).
- **xBSD**: Die derzeit noch beste und vollständigste Unterstützung für IPv6 - vor allem ein Verdienst des japanischen KAME-Projektes.
- **Linux**: Der Kernel 2.4 bietet eine als "experimental" markierte Unterstützung für IPv6, der jedoch noch wichtige Features wie IPsec und Privacy Extensions (RFC 3041) fehlen. Der neue Kernel 2.6.x bietet eine umfassende IPv6-Unterstützung auf ähnlichem Niveau wie die BSD-Derivate.
- **Mac OS X**: Enthält seit Version 10.2 Unterstützung für IPv6 auf der Basis von KAME. Erst seit Version 10.3 lässt sich IPv6 auch über das GUI konfigurieren.
- **Cisco**: Cisco ist der Marktführer bei Internetroutern. Beim Deployment von IPv6 gehörte Cisco aber eher zu den Schlusslichtern. Es gibt zwar mittlerweile IPv6-fähige IOS-Versionen, diese werden aber nicht standardmässig installiert, da sie noch als experimentell eingestuft werden.

Viele Anwendungen (vor allem aus dem Bereich der Freien Software) sind inzwischen ebenfalls IPv6-fähig. Im heimatlichen LAN kann man so schon relativ problemfrei IPv6 benutzen. Jenseits des eigenen Border-Routers sieht es derzeit noch düster aus: Es gibt keine Provider, die native IPv6-Anbindung verkaufen, so dass man im Moment auf bisweilen unbefriedigende Bastel-Lösungen mit Tunneling zurückgreifen muss.

Zumindest in Europa und Amerika besteht einfach auch keine Notwendigkeit zur Migration zu IPv6, da noch genügend freie IPv4-Adressen vorhanden sind (in Europa über 50%). In Asien geht der Trend inzwischen dahin, bei Neubauten (z. B. dem NTT-Backbone) IPv6 auch zu benutzen. Von seiten der Endbenutzer wird IPv6 auch deshalb nicht gefordert, weil außer dem größeren Adressbereich die wesentlichen neuen Features von IPv6 inzwischen mehr oder weniger erfolgreich nach IPv4 rückportiert wurden (z. B. IPsec, QoS, Multicast. Das Renumbering und die Autokonfiguration kann man durch DHCP angenähert erreichen) - es gibt keine "Killer-Applikation", die nur mit IPv6 funktionieren würde.

In Deutschland federführend bei den Versuchen zu IPv6 ist das JOIN-Projekt der Uni Münster. Im Moment ist seitens des JOIN und des DFN ein erstes IPv6-Backbone in Deutschland im Aufbau. Die Planungen für das "6Win" sehen einen ringförmigen Backbone durch Deutschland mit Querverbindung zwischen Essen und Berlin vor. Parallel dazu baut die Deutsche Telekom einen eigenen IPv6-Backbone zwischen den Standorten Darmstadt, Münster und Berlin auf und bietet ihren Geschäftskunden im Rahmen eines Showcase-Projektes Anschluss daran an. Dieses Netz soll in Münster und Berlin mit dem 6Win verbunden werden. Ebenfalls in Münster liegt der deutsche zentrale Zugang zum experimentellen IPv6-Netzwerk 6Bone.

Was ist mit IPv5?

Wenn es IPv4 und IPv6 gibt, was ist dann mit IPv5? gehört mit zu den häufigsten Fragen, die sich IPv6-Neueinsteiger stellen. Ein Protokoll mit dem Namen IPv5 gibt es nicht, allerdings hat die IANA die IP-Versionsnummer 5 für das Internet Stream Protocol Version 2 (ST2, definiert in RFC 1819) reserviert, das gegenüber IPv4 verbesserte Echtzeitfähigkeiten haben sollte, dessen Entwicklung dann aber zu Gunsten von IPv6 und RSVP eingestellt wurde.

UDP

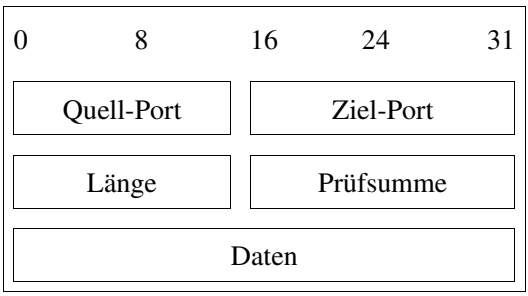
Das *User Datagram Protocol* (**UDP**) ist ein minimales, verbindungsloses Netzwerkprotokoll. Es gehört zur Transportschicht der TCP/IP-Protokollfamilie und ist im Gegensatz zu TCP nicht auf Zuverlässigkeit ausgelegt.

Anwendung	DNS	DHCP	NTP	...
Transport	**UDP**			
Netzwerk	IP			
Netzzugang	Ethernet	Token Ring	FDDI	...

UDP im TCP/IP-Protokollstapel

Header-Format

Der UDP-Header besteht aus 4 Headerfeldern, von denen zwei optional sind. Die Quell- und Ziel-Port Felder sind 16 Bit groß und identifizieren den sendenen und empfangenden Prozess. Da UDP verbindungslos ist, ist der Quell-Port optional. Er wird dann 0 gesetzt. Den Portfeldern folgt das verbindliche Längenfeld, das die Größe der Daten des UDP-Datagramms in Oktetten enthält. Der kleinstmögliche Wert ist 8 Oktette. Das letzte Headerfeld ist eine 16 Bit große Prüfsumme über den Header und die Daten. Die Prüfsumme ist auch optional, wird aber in der Praxis fast immer benutzt (falls nicht, wird sie ebenfalls 0 gesetzt). Dem Header folgen anschließend die Daten.



Eigenschaften

Verbindungslos heißt, es wird nicht erst eine Verbindung zum Gegenüber aufgebaut (Handshaking wie bei TCP), sondern man schickt "auf gut Glück" eine Anfrage. Es ist also nicht garantiert, dass das Paket überhaupt ankommt.

Aufgrund dieser Tatsache können zwischen zwei Hosts relativ schnell Datenpakete ausgetauscht werden. Es wird deshalb dort eingesetzt, wo die schnelle Übermittlung wichtiger ist als die Zuverlässigkeit, also die Gewissheit, dass die Daten korrekt und vollständig angekommen sind. In der Praxis sind das Übertragungen von Multimedia oder bei Online-Spielen. Auch ein sehr wichtiger Dienst im Internet, das Domain Name System, setzt auf UDP auf.

Zeitlicher Versatz der Pakete (engl. jitter) kann bei UDP nicht erkannt werden.

Verwendung

UDP wird unter anderem von folgenden Protokollen verwendet:

- DNS (Domain Name System)
- NFS (Network File System)
- TFTP (Trivial File Transport Protocol)
- SNMP (Simple Network Management Protocol)

UDP selber verwendet meistens das Internet Protocol.

Das Protokoll findet man im RFC 768 - User Datagram Protocol (*http://www.ietf.org/rfc/rfc768.txt*)

Transmission Control Protocol

Das *Transmission Control Protocol* (**TCP**) ist ein zuverlässiges, verbindungsorientiertes Transportprotokoll in Computernetzwerken. Es ist Teil der TCP/IP-Protokollfamilie.

TCP stellt einen virtuellen Kanal zwischen zwei Rechnern (genauer: Endpunkten) her. Auf diesem Kanal können in beide Richtungen Daten übertragen werden. TCP setzt in den meisten Fällen auf das IP-Protokoll auf. Es ist in Schicht 4 des OSI-Netzwerkschichtenmodells angesiedelt.

Anwendung	FTP	SMTP	HTTP	...
Transport	**TCP**			
Netzwerk	IP			
Netzzugang	Ethernet	Token Ring	FDDI	...

TCP im TCP/IP-Protokollstapel

Verbindungsauf- und -abbau

Ein Endpunkt stellt ein Tupel bestehend aus IP-Adresse und Port dar. Ports sind 16-bit Zahlen und reichen von 1 bis 65535. Ports unterhalb von 1024 sind reserviert (englisch: *well known ports*) und werden von der IANA vergeben, z. B. ist Port 25 für das SMTP Protokoll für elektronische Post reserviert.

Jede TCP-Verbindung wird eindeutig durch zwei Endpunkte definiert. So ist es möglich, dass ein Webserver auf Port 80 mehr als eine Verbindung zu einem anderen Rechner geöffnet haben kann.

Ein Server-Rechner, der einen Dienst wie beispielsweise elektronische Post anbietet, generiert einen Endpunkt mit dem Port und seiner Adresse (er kann auch beliebige Adressen zulassen). Dies wird als *passive open* bezeichnet.

Will ein Client eine Verbindung aufbauen, generiert er einen eigenen Endpunkt aus seiner Rechneradresse und einer noch freien Portnummer. Mit Hilfe eines ihm bekannten Ports und der Adresse des Servers kann dann eine Verbindung aufgebaut werden. Für den Aufbau der Verbindung sind drei Pakete erforderlich (3-Wege-*Handshake*).

Während der Datenübertragungsphase (*active open*) sind die Rollen von Client und Server (aus **TCP**-Sicht) vollkommen symmetrisch. Insbesondere kann jeder der beiden beteiligten Rechner einen Verbindungsabbau einleiten. Während des Abbaus kann die Gegenseite noch Daten übertragen, die Verbindung kann also *halb*-offen sein. Ein 4-Wege-*Handshake* wird benutzt, um die Verbindung abzubauen.

Zuverlässigkeit

Im Gegensatz zum paketorientierten UDP implementiert **TCP** einen bidirektionalen, byteorientierten, zuverlässigen Datenstrom zwischen zwei Endpunkten. Das darunterliegende Protokoll (meist IP) ist paketorientiert, wobei Datenpakete verlorengehen können, in verkehrter Reihenfolge ankommen dürfen und sogar doppelt empfangen werden können. TCP prüft die Integrität der Daten mittels einer Prüfsumme und stellt die Reihenfolge durch Sequenznummern sicher. Der Sender wiederholt das Senden von Paketen falls keine Bestätigung innerhalb einer bestimmten Zeitspanne (Timeout) eintrifft. Die Daten der Pakete werden im Empfänger in einem Puffer zu einem Datenstrom zusammengefügt und doppelte Pakete verworfen.

Bestätigungen

Die jeweilige Länge des Puffers, bis zu der keine Lücke im Datenstrom existiert, wird bestätigt (*Windowing*). Dadurch ist die Ausnutzung der Netzwerk-Bandbreite auch bei großen Strecken möglich. Bei einer Übersee- oder Satellitenverbindung dauert das Eintreffen des ersten *Acknowledges* (ACK) aus technischen Gründen mehrere 100 ms, in dieser Zeit können unter Umständen mehrere hundert Pakete gesendet werden. Der Sender kann den Empfängerpuffer füllen bevor die erste Bestätigung eintrifft. Alle Pakete im Puffer können gemeinsam bestätigt werden. Bestätigungen werden zusätzlich zu den Daten in die Paket-Header im entgegengesetzen Datenstrom eingefügt (*Piggybacking*).

Weitere Protokolleigenschaften

Über ein Dringlichkeitsbit (*Urgent*) können Daten als vorrangig gekennzeichnet werden. Dadurch ist beispielsweise die bevorzugte Behandlung von CTRL-C (Abbruch) bei einer Terminalverbindung (TELNET) möglich.

Um Bandbreite zu sparen, wird auf der TCP Ebene meistens der Nagle-Algorithmus eingesetzt.

Datenwiederholung

Die Wiederholung von Daten für die noch keine Bestätigung empfangen wurde, ist nicht unproblematisch. Im Internet, in dem viele Netzwerke mit unterschiedlichen Eigenschaften verbunden werden, ist Datenverlust einzelner Pakete durchaus normal. Die Verlustrate nimmt zu, falls bestimmte Netzwerke innerhalb der Verbindung an ihre Auslastungsgrenze kommen. Eine naive Implementierung von TCP/IP würde die verlorenen Pakete einfach wiederholen, was zu noch größerer Auslastung führen würde und unter Umständen zum Zusammenbruch des Netzwerks führen könnte. TCP/IP Implementierungen verwenden daher Algorithmen, die dies verhindern. Normalerweise wird langsam gestartet (*Slow Start*) und die Senderate dann bis zum Datenverlust erhöht. Jeder Datenverlust verringert die Senderate, insgesamt nähert sich die Datenrate dem jeweiligen zur Verfügung stehenden Maximum.

HTTP

Das **Hypertext Transfer Protocol** (HTTP) ist ein zustandsloses Protokoll in der Anwendungsschicht. Es dient zur Übertragung von Hypermedia-Informationen. Durch Erweiterung seiner Anfragemethoden, Headerinformationen und Fehlercodes ist es nicht auf Hypertext beschränkt. Es wird von Web-Browsern verwendet um auf Web-Server zuzugreifen.

Anwendung	**HTTP**			
Transport	TCP			
Netzwerk	IP			
Netzzugang	Ethernet	Token Ring	FDDI	...

HTTP im TCP/IP-Protokollstapel

Das Protokoll wurde 1989 von Tim Berners-Lee am CERN zusammen mit dem URL und HTML entwickelt: das World Wide Web (WWW) wurde geboren.

HTTP ist ein Kommunikationsschema, um Webseiten (oder Bilder oder prinzipiell jede andere beliebige Datei) von einem entfernten Computer auf den eigenen zu übertragen. Wenn auf einer Webseite der Link *www.example.net:80/infotext.html* angeklickt wird, so wird an den Computer mit dem Namen *www.example.net* die Anfrage gerichtet, die Datei *infotext.html* zurückzusenden. Der Name *www.example.net* wird dabei zuerst über das DNS-Protokoll in eine Adresse umgesetzt. Zur Übertragung wird über das TCP-Protokoll auf Port 80 eine HTTP-GET Anforderung gesendet.

Zusätzliche Informationen wie Angaben über den Browser, gewünschte Sprache etc. können über einen Header in jeder HTTP-Kommunikation übertragen werden. Der Computer, der einen Web-Server (an Port 80) betreibt, sendet dann seinerseits eine HTTP-Antwort zurück. Diese besteht aus Headerinformationen des Servers, einer Leerzeile und dem Inhalt der Datei *infotext.html*. Die Datei ist normalerweise im Hypertext-Format HTML, das vom Browser in eine lesbare und ansprechende Darstellung gebracht wird. Es kann jedoch jede andere Datei in jedem beliebigen Format sein, zum Beispiel Bildinformationen, Audio- und Videodateien.

Anfrage:

```
GET /infotext.html HTTP/1.1
Host: www.example.net:80
```

Antwort:

```
HTTP/1.1 200 OK
Server: Apache/1.3.29 (Unix) PHP/4.3.4
Content-Length: (Größe von infotext.html)
Content-Language: de
Content-Type: text/html
Connection: close
(Inhalt von infotext.html)
```

Die Information kann auch dynamisch generiert werden und braucht auf dem Server nicht als Datei abgelegt sein. Der Server sendet eine Fehlermeldung zurück, wenn die Information aus irgendeinem Grund nicht gesendet werden kann. Der genaue Ablauf dieses Vorgangs (Anfrage und Antwort) ist in der HTTP-Spezifikation festgelegt.

Bei HTTP 1.0 wird vor jeder Anfrage eine neue TCP-Verbindung aufgebaut und nach Übertragung der Antwort wieder geschlossen. Enthält eine HTML-Datei Verweise auf zehn Bilder, so werden insgesamt elf TCP-Verbindungen benötigt, um die Seite auf einem grafik-

fähigen Browser aufzubauen. In der neuesten Version 1.1 von HTTP, können mehrere Anfragen pro TCP-Verbindung gemacht werden. Für die HTML-Datei mit zehn Bildern, wird so nur eine TCP-Verbindung benötigt. Zusätzlich können abgebrochene Downloads fortgesetzet werden und eine Menge auf der Low-Level-Ebene wurde verbessert. Informationen aus früheren Anforderungen gehen verloren (zustandsloses Protokoll). Über Cookies in den Headerinformationen können aber Anwendungen realisiert werden, die Statusinformationen (Benutzereinträge, Warenkörbe) zuordnen können. Dadurch können Anwendungen die Status- bzw. Sitzungseigenschaften erfordern, realisiert werden. Auch eine Benutzerauthentifizierung ist möglich. Normalerweise kann die Information, die über HTTP übertragen wird, auf allen Rechnern und Routern, die im Netzwerk durchlaufen werden, gelesen werden. Über HTTPS kann die Übertragung verschlüsselt erfolgen.

HTTP-Statuscodes

1xx: Informationen

- 100: Continue
- 101: Switching Protocols

2xx: Erfolgreiche Operation

- 200: OK
- 201: Created
- 202: Accepted
- 203: Non-Authoritative Information
- 204: No Content
- 205: Reset Content
- 206: Partial Content

3xx: Umleitung

- 300: Multiple Choices
- 301: Moved Permanently
- 302: Found
- 303: See Other
- 304: Not Modified
- 305: Use Proxy
- 307: Temporary Redirect

4xx: Client-Fehler

- 400: Bad Request
- 401: Unauthorized
- 402: Payment Required
- 403: Forbidden
- 404: Not Found
- 405: Method Not Allowed
- 406: Not Acceptable
- 407: Proxy Authentication Required
- 408: Request Time-out
- 409: Conflict
- 410: Gone
- 411: Length Required
- 412: Precondition Failed
- 413: Request Entity Too Large
- 414: Request-URI Too Large
- 415: Unsupported Media Type
- 416: Requested range not satisfiable
- 417: Expectation Failed

5xx: Server-Fehler:

- 500: Internal Server Error
- 501: Not Implemented
- 502: Bad Gateway
- 503: Service Unavailable
- 504: Gateway Time-out
- 505: HTTP Version not supported

Das Protokoll ist in folgenden RFCs definiert:

- RFC 2616: Hypertext Transfer Protocol - HTTP/1.1
- RFC 2617: HTTP Authentication: Basic and Digest Access Authentication

FTP

FTP steht für **F**ile **T**ransfer **P**rotocol, was im Deutschen in etwa *Dateiübertragungsverfahren* bedeutet.

Dieses Netzwerk-Protokoll wird benutzt, um Dateien in TCP/IP-Netzwerken zwischen einem Server (hier speziell: einem FTP Server) und einem FTP Client auszutauschen (Download (Server–>Client). Das Protokoll kann aber auch zum Upload (Client–>Server) benutzt werden. Es ist in RFC 959 definiert. Im Gegensatz zu dem Protokoll sftp werden alle Daten, also auch das Kennwort, unverschlüsselt übertragen.

FTP benutzt zwei Ports, nämlich den *data*-Port 20 zur Datenübertragung und den *command*- oder *control*-Port 21 zur Befehlsübertragung. Es kennt zwei unterschiedliche Übertragungsmodi. In beiden Fällen initiiert der Client auf Port 21 des FTP-Servers die Verbindung:

1. Beim **Aktiv Mode** beginnt der Server auf Port 20 eine Datenverbindung zum Client (Port > 1023) hin aufzubauen.
2. Beim **Passiv Mode** beginnt der Client eine Datenverbindung zum Server hin aufzubauen. Auf beiden Seiten wird ein Port > 1023 benutzt. Der Port 20 wird in diesem Fall nicht gebraucht.

Der Vorteil des weitverbreiteten und standardisierten FTP-Protokolls ist, dass Computer trotz komplett verschiedener Betriebssysteme, wie etwa Unix und Windows, untereinander Dateien austauschen können. Ein Nachteil ist, dass Passwörter und sonstige Daten unverschlüsselt, also in Klartext übermittelt werden.

Viele FTP-Server erlauben so genanntes *"anonymous ftp"*. Das heißt, Benutzer brauchen kein Benutzerkonto auf dem Server, sondern sie können sich durch einfaches Eingeben ihrer E-Mail-Adresse als Passwort auf dem Server einloggen.

Anwendung	**FTP**			
Transport	TCP			
Netzwerk	IP			
Netzzugang	Ethernet	Token Ring	FDDI	...

FTP im TCP/IP-Protokollstapel

Beispiele für FTP-Clients

- Unter Windows: FileZilla, WS-FTP, LeechFTP

- Unter Unix: gftp, kbear , AxY FTP
- Unter Mac OSX: Fetch, Captain FTP, Anarchie

Fast alle Webbrowser haben ebenfalls FTP implementiert und können als Clients ebensolcher benutzt werden. Beispiel-URL für FTP: ftp://ftp.example.org/

Der Benutzername und das Passwort können auch direkt in die URL eingebaut werden: ftp://login:password@ftp.example.org

Die gebräuchlichsten ftp-Kommandos

Die gebräuchlichsten Kommandos des Kommandozeilenprogramms "ftp", welches FTP auf der Client-Seite versteht. (Bei einem grafischen FTP Programm funktioniert die Bedienung auf Klick.)

Befehl	**Beschreibung**
open	Öffnet eine Verbindung zum Server
user	Definiert den User, der sich einloggen will. Fast immer kann hier **anonymous** angegeben werden, um eine anonyme Verbindung herzustellen. Da der User in diesem Fall dem Server nicht bekannt ist, hat er normalerweise auch nicht all Rechte und kann zum Beispiel nur auf bestimmte Verzeichnisse zugreifen. Diese Art der Verbindung wird als **Anonymous-ftp** bezeichnet.
close	Beendet eine Verbindung zum Server.
cd	Wechselt in ein anderes Verzeichnis auf dem ftp-Server. Statt "\" wie unter DOS und Windows wird als Trennzeichen der normale Schrägstrich "/" (Unix-Konvention) verwendet.
list	Fordert das Ergebnis eines Verzeichnisdurchsuchung auf. Es können auch das zu durchsuchende Verzeichnis und ein Suchmuster angegeben werden.
get	Kopiert eine Datei vom Server auf den lokalen Rechner. Das Kopieren funktioniert aber nur, wenn man Leserechte auf dem Server hat.
put	Kopiert eine Datei vom lokalen Rechner auf den Server. Funktioniert aber nur, wenn man Schreibrechte auf dem Server hat. Viele Server stellen für diesen Zweck ein incoming-Verzeichnis zur Verfügung, in das Dateien abgelegt werden dürfen.
mget	Kopiert eine oder mehrere Dateien vom Server auf den lokalen Rechner. Vor jedem Kopiervorgang wird die Datei angezeigt und der Anwender gefragt, ob diese übertragen werden soll. Dies kann man mit dem ftp-Kommando "prompt" ab- und anschalten. Das Kopieren funktioniert aber nur, wenn man Leserechte auf dem Server hat.

Befehl	Beschreibung
mput	Kopiert eine oder mehrere Dateien vom lokalen Rechner auf den Server. Vor jedem Kopiervorgang wird die Datei angezeigt und der Anwender gefragt, ob diese übertragen werden soll. Dies kann man mit dem ftp-Kommando "prompt" ab- und anschalten. Das Kopieren funktioniert aber nur, wenn man Schreibrechte auf dem Server hat. Viele Server stellen für diesen Zweck ein incoming-Verzeichnis zur Verfügung, in das Dateien abgelegt werden dürfen.
prompt	Damit kann man die Rückfragen bei dem Kommandos mget und mput an- oder abschalten.
mkdir	Erzeugt ein neues Verzeichnis auf dem Server
delete	Löscht eine Datei auf dem Server.
mdelete	Löscht mehrere Dateien über eine Maske
rename	Mit dieser Funktion können Dateien und Verzeichnisse auf dem Server umbenannt werden.

TELNET

Telnet ist der Name eines im Internet weit verbreiteten Protokolls. Das IETF Dokument STD 8, in welchem es beschrieben wurde, beginnt folgendermaßen: "Der Sinn des TELNET Protokolls besteht darin eine ziemlich allgemeine, bi-direktionale, 8-bit pro Byte orientierte Kommunikationsmöglichkeit zu bieten." Es wird üblicherweise dazu verwendet Benutzern den Zugang zu Internetrechnern über die Kommandozeile zu bieten. Sicherheitsexperten lehnen dieses Protokoll inzwischen ab.

Weiterhin ist telnet der Name eines Programmes, welches Telnetverbindungen zu einem entfernten Gastrechner ermöglicht. Das Telnetprogramm stellt dabei die benötigten Client Funktionen des Protokolls zur Verfügung. Manchmal wird das Wort auch in Verbform verwendet um die Tätigkeit des Verbindens zu einem entfernten Rechner (Server) zu beschreiben, wie in "Wenn du dein Passwort ändern willst, dann mach einen telnet auf den Server und rufe passwd auf".

Anwendung	**Telnet**			
Transport	TCP			
Netzwerk	IP			
Netzzugang	Ethernet	Token Ring	FDDI	...

Telnet im TCP/IP-Protokollstapel

Telnet ist ein Client-Server-Protokoll, es verwendet TCP und die Clients verbinden sich meistens über Port 23 mit dem Zielrechner (allerdings lässt sich dieser Port wie bei den meisten Internetprotokollen auch ändern). Teilweise aufgrund des Protokolldesigns, teilweise aufgrund der üblicherweise zur Verfügung gestellten Flexibilität der Telnetprogramme, ist es ebenfalls möglich mit einem Telnetprogramm eine interaktive TCP Verbindung zu einigen anderen Internetservices aufzubauen. Eine klassische Verwendung davon ist beispielsweise eine Telnetverbindung über Port 25 aufzubauen (wo sich meistens ein SMTP Server befindet) um Fehler in einem eMailserver zu finden.

Für Netzwerkadministratoren ist telnet ein nützliches Tool. Damit kann man feststellen, ob eine TCP/IP-Verbindung über einen bestimmten Port zustande kommt oder ob dieser Port evtl. von einer Firewall geblockt ist.

Das Telnetprotoll kann man aufteilen in den Kernbereich, sowie einem Satz von Erweiterungen. Das Kernprotokoll wird in den IETF Dokumenten RFC 854 und RFC 855 (gesammelt in STD 8) beschrieben. STD 8 beschreibt einige grundsätzliche Arbeitsweisen des Protokolls und die Möglichkeiten Erweiterungen zu definieren und zu implementieren. Es gibt zahlreiche Erweiterungen, einige wurden als Internetstandards aufgenommen, andere nicht. Die IETF STD Dokumente 27-32 definieren verschiedene Telneterweiterungen (die meistens implementiert und verwendet werden). Von den übrigen Erweiterungen sind die nützlichsten vermutlich diejenigen, welche bereits als IETF Standard vorgeschlagen wurden (der übliche Weg zur Standardisierung); Details können im Dokument STD 1 gefunden werden.

Sicherheit

Zusätzlich zu den vorhandenen Sicherheitslücken in Telnetservern, gibt es noch das Problem, dass Daten (inklusive Einwahlpasswörtern) nicht verschlüsselt werden. Daher kann jeder der dazu in der Lage ist die Verbindungsdaten einzusehen und somit auf einfache Weise vertrauliche Daten mitlesen und Zugriff auf die Verzeichnisse des Benutzers auf dessen Zielrechner erlangen. Aufgrund dieser Mängel wird der Einsatz des Telnetprotkolls derzeitig stark reduziert und stattdessen das sicherere und funktionalere Protokoll SSH, welches 1998 veröffentlicht wurde, verwendet. Computersicherheitsexperten wie SANS und Mitglieder von comp.os.linux.security (eine Newsgroup) empfehlen den Einsatz von Telnet für remote logins unter allen Umständen gänzlich einzustellen.

Secure Shell

Secure shell oder **SSH** ist sowohl ein Programm als auch ein Netzwerkprotokoll, mit dessen Hilfe man sich z. B. über das Internet auf einem entfernten Computer einloggen und dort Programme ausführen kann. Die IANA hat dem Protokoll den Port 22/TCP zugeordnet.

Anwendung	SSH			
Transport	TCP			
Netzwerk	IP			
Netzzugang	Ethernet	Token Ring	FDDI	...

SSH im TCP/IP-Protokollstapel

SSH wurde von **Tatu Ylönen** entwickelt. Es ermöglicht eine sichere authentifizierte und verschlüsselte Verbindung zwischen zwei Rechnern über ein unsicheres Netzwerk. Deswegen sollten rlogin, telnet und rsh nicht mehr verwendet, sondern stattdessen ssh eingesetzt werden. X11-Sitzungen und andere TCP/IP-Verbindungen können ebenfalls über diesen sicheren Kanal weitergeleitet werden.

SSH ist ursprünglich ein Unix-Programm, es gibt aber inzwischen auch Portierungen auf andere Plattformen wie beispielsweise Microsoft Windows (PuTTY ist hier inzwischen sehr populär).

Eine neuere Version des Protokolls wurde unter dem Namen SSH2 herausgegeben. SSH2 zeichnet sich durch einen modularen Aufbau der Transport-, Autorisierungs- und Verbindungsschichten aus und ermöglicht im Gegensatz zu SSH1 die Verwendung von verschiedenen Verschlüsselungsalgorithmen.

Die von SSH1 verwendete Integritätsprüfung weist Schwachstellen auf, die es einem Angreifer ermöglichen, eine SSH1-Sitzung auszuspähen. Daher sollte nur noch die neue Protokollversion SSH2 verwendet werden.

Mit OpenSSH existiert auch eine freie Implementierung von SSH.

Eine Arbeitsgruppe der Internet Engineering Task Force (IETF) namens secsh arbeitet momentan an der Standardisierung des Protokolles.

Domain Name System

Das **Domain Name System** (DNS) ist einer der wichtigsten Dienste im Internet. Das DNS ist eine verteilte Datenbank, die den Namensraum im Internet verwaltet. Das geschieht im Wesentlichen durch die Umsetzung von Namen in Adressen. Das Ganze ist vergleichbar mit einem Telefonbuch, das die Namen der Teilnehmer in ihre Telefonnummer auflöst.

Das DNS ist notwendig, weil Menschen sich Namen weitaus einfacher merken können als Zahlenkolonnen. So kann man sich den Domainnamen *www.wikipedia.de* sehr einfach merken, die dazugehörende IP-Adresse *81.2.181.51* dagegen nicht ganz so einfach. Darüber hinaus ermöglicht das DNS eine logische Lösung von der darunterliegenden Physik, z. B. Änderung der IP-Adresse, ohne den Domainnamen ändern zu müssen, und sogar rudimentäre Lastverteilung.

Anwendung	DNS			
Transport	UDP		TCP	
Netzwerk	IP			
Netzzugang	Ethernet	Token Ring	FDDI	...

DNS im TCP/IP-Protokollstapel

Das DNS wurde 1983 von Paul Mockapetris entworfen und im RFC 882 beschrieben. Der RFC 882 wurde inzwischen von den RFCs 1034 und 1035 abgelöst.

Das DNS löste die *hosts*-Dateien ab, die bis dahin für die Namensauflösung zuständig waren. Das DNS zeichnet sich aus durch

- dezentrale Verwaltung
- hierarchische Strukturierung des Namensraums in Baumform
- Eindeutigkeit der Namen
- Erweiterbarkeit

Komponenten des DNS

Das DNS besteht aus drei Hauptkomponenten

- Domänennamensraum und die Resource Records (RR)
- Nameservern
- Resolver

Domänennamensraum

Der Domänennamensraum hat eine baumförmige Struktur. Die Blätter und Knoten des Baumes werden als **Labels** bezeichnet. Ein kompletter Domänenname eines Objektes besteht aus der Verkettung aller Labels. Label sind Zeichenketten (alphanumerisch, als einziges Sonderzeichen ist der '-' erlaubt) die mindestens ein Zeichen und maximal 63 Zeichen lang sind. Die einzelnen Label werden durch Punkte voneinander getrennt. Ein Domänenname wird mit einem Punkt abgeschlossen (Der hinterste Punkt wird normalerweise weggelassen, gehört rein formal aber zu einem vollständigen Domänennamen dazu). Ein korrekter, vollständiger Domänenname lautet z. B. *www.wikipedia.de.* (der Punkt gehört zum Domänennamen).

Ein Domänenname darf inklusive aller Punkte maximal 255 Zeichen lang sein.

Ein Domänenname wird immer von rechts nach links delegiert und aufgelöst, d. h. je weiter rechts ein Label steht, umso höher steht es im Baum. Der Punkt ganz rechts wird auch als *root* (Wurzel) im Namensraum bezeichnet.

Das erste Label (das, das ganz rechts steht) wird im allgemeinen auch als Top Level Domain (TLD) bezeichnet.

Resource Records (RR)

Die DNS-Objekte einer Domäne (zum Beispiel die Rechnernamen) werden als Satz von Resource Records in einer **Zonendatei** gehalten, die auf einem oder mehreren autoritativen Nameservern abliegt. Anstelle von *Zonendatei* wird meist der etwas allgemeinere Ausdruck **Zone** verwendet.

Ein Resource Record ist immer folgendermaßen aufgebaut:

```
<name> [<ttl>] [<class>] <type> <rdata>
```

- **<name>** Der Domänenname des Objekts zu dem der Resource Record gehört
- **<ttl>** *time to live* Gültigkeit des Resource Records (in Sekunden)
- **<class>** Protokollgruppe zu der der Resource Record gehört. Üblicherweise wird **IN** (Internet) verwendet. Es sind aber auch die Klassen **CH** (Chaosnet) oder **HS** (Hesiod) sowie **CS** (CSNET, wird nicht mehr verwendet und wird lediglich noch als Beispiel in einigen obsoleten RFCs genannt) möglich.
- **<type>** beschreibt den Typ des Resource Records. Im DNS mögliche Typen sind:
  - **A** IP-Adresse eines Hosts
  - **CNAME** Alias name für einen Host
  - **HINFO** Host information
  - **KEY** enthält einen dem Namen zugeordneten Public Key
  - **MB** Mailbox domain name (*Experimentell*)
  - **MD** Mail destination (nicht mehr in Gebrauch - heutzutage wird MX verwendet)
  - **MF** Mail forwarder (nicht mehr in Gebrauch - heutzutage wird MX verwendet)
  - **MG** Mail group member (*Experimentell*)
  - **MINFO** Mailbox oder *mail list information*
  - **MR** Mail rename domain name (*Experimentell*)
  - **MX** Mail Exchange – der für diese Domain zuständige Mailserver
  - **NULL** Null Resource Record (*Experimentell*)
  - **NS** Hostname eines autoritativen Nameservers
  - **PTR** Domain Name Pointer (für das Reverse Mapping, um IP Adressen Namen zuzuweisen)
  - **SIG** enthält eine digitale Unterschrift (wird von DNSSEC (=*DNS Security*) verwendet)
  - **SOA** Start of Authority
  - **TXT** Text

- **WKS** *Well known service description*

- **<rdata>** (resource data) Daten die den Resource Record näher beschreiben (z. B. eine IP Adresse für einen A-RR, oder einen Hostnamen für einen NS-RR)

Nameserver

Nameserver sind Programme die einen oder mehrere Teile des Namensraumes autoritativ kennen, und diese auf Anfrage weitergeben. Nameserver werden von der höheren Ebene im Baum delegiert und sind dann für den Teilnamensraum unterhalb der delegierten Ebene zuständig (und können u.U. weitere Teilnamensräume unterhalb dieser Ebene delegieren). Die Baumstruktur stellt die eindeutige Zuordnung eines Nameservers zu einem Teil des Namensraum sicher.

Normale Anfragen werden auf Port 53 UDP beantwortet. Transfers kompletter Zonen werden auf Port 53 TCP durchgeführt.

Früher sprach man von *primary* und *secondary* Nameserver, heute spricht man von **autoritativen** Nameservern. Ein autoritativer Nameserver ist ein Nameserver, der gesicherte Informationen über eine Zone hat. Dem gegenüber steht ein **nicht-autoritativer** Nameserver der Informationen über eine Zone sozusagen aus zweiter oder dritter Hand hat, also nicht sicher sagen kann, dass die Information korrekt ist (da sie sich z. B. schon geändert haben kann).

Nameserver speichern die einmal von einem Resolver angefragten Informationen im lokalen RAM ab, damit diese bei einer erneuten Anfrage schneller vorliegen. Diese Technik wird als **Caching** bezeichnet. Caching ist möglich, da sich DNS-Daten normalerweise nur sehr selten ändern. Die Daten im Cache des Nameservers verfallen nach der TTL (*time to live*). Das kann u.U. aber auch bedeuten, dass der Nameserver in dieser Zeit falsche Informationen liefern kann, wenn sich die Daten zwischenzeitlich geändert haben. Nameserver können auch als **caching only Nameserver** agieren. Sie besitzen dann selbst keine gesicherten (autoritative) Informationen sondern lösen alle eintreffenden Anfragen rekursiv auf.

Damit ein Nameserver Informationen über andere Teile des Namensraumes finden kann, werden ihm Informationen über die sogenannte Root-Server in Form einer statischen Datei hinterlegt. Diese *Cachedateien* enthalten die Namen und IP Adressen der Root-Server. Derzeit gibt es 13 Root-Server (Server A bis M).

Nameserversoftware

- BIND (Berkeley Internet Name Domain) ist der Ur-Nameserver und heute noch die meistgenutzte Nameserversoftware. BIND ist Open Source Software.
- djbdns (entwickelt von Dan Bernstein) gilt als sehr sicher und erfreut sich steigender Beliebtheit.
- PowerDNS ist eine Implementierung, die vor allem für das direkte Betreiben von Zonen aus SQL-Datenbanken bekannt ist.
- NSD ist optimiert für Server die ausschließlich autoritative Antworten liefern sollen.

Resolver

Resolver sind Programme die Informationen aus den Nameservern abrufen können. Sie bilden die Schnittstelle zum Nameserverdienst. Resolver sind entweder eigene Programme, oder sie sind in Applikationen (z. B. einen Browser) eingebunden. Ein Resolver arbeitet entweder **iterativ** oder **rekursiv**.

Bei einer rekursiven Anfrage schickt der Resolver eine Anfrage an einen ihm bekannten Nameserver und gibt als Antwort entweder den gewünschten Resource Record (wenn der befragte Nameserver selber rekursiv arbeitet) oder "gibt es nicht". Rekursiv arbeitende Resolver überlassen also die Arbeit anderen und funktionieren so wie manches andere im Internet: *Ich weiß ein bisschen was und ich kenne jemanden der mehr weiß.*

Bei einer iterativen Anfrage bekommt der Resolver entweder den gewünschten Resource Record oder einen weiteren Nameserver, den er als nächsten fragt. Der Resolver hangelt sich so von Nameserver zu Nameserver bis er bei einem autoritativen Nameserver landet. Die so gewonnene Antwort gibt der Resolver dann weiter.

Die Root-Server arbeiten ausschließlich iterativ. Sie wären sonst mit der Anzahl der Anfragen schlicht überlastet.

Bekannte Resolver sind die Programme *nslookup* und *dig*.

Erweiterung des DNS

Bisher waren die Label – wie beschrieben – auf alphanumerische Zeichen und das Zeichen '-' eingeschränkt. Dies hängt vor allem damit zusammen, dass das DNS (wie auch das Internet ursprünglich) in den USA entwickelt wurde. Allerdings gibt es in vielen Ländern Zeichen, die nicht in einem Label verwendet werden durften (in Deutschland zum Beispiel die Umlaute) oder Zeichen aus komplett anderen Schriftsystemen (z. B. Chinesisch). Namen mit diesen Zeichen waren bisher nicht möglich.

Dies hat sich durch die Einführung von IDNA (RFC 3490) geändern. Seit März 2004 können deutsche, lichtensteinische, österreichische und schweizer Domains mit Umlauten registriert und verwendet werden. Um das neue System mit dem bisherigen kompatibel zu halten, werden die erweiterten Zeichensätze mit erlaubten Zeichen kodiert, also auf derzeit gültige Namen abgebildet. Die erweiterten Zeichensätze werden dabei zunächst gemäß dem *Nameprep*-Algorithmus (RFC 3491) normalisiert, und anschließend per *Punycode* (RFC 3492) auf den für DNS verwendbaren Zeichensatz abgebildet. Das Voransetzen des durch die IANA festgelegten IDNA-Prefix xn-- vor das Ergebnis der Kodierung ergibt das vollständige IDN-Label.

Eine weitere aktuelle Erweiterung des DNS stellt ENUM (RFC 2916) dar. Diese Anwendung ermöglicht die Addressierung von Internet-Diensten über Telefonnummern, also das "Anwählen" von per Internet erreichbaren Geräten mit dem aus dem Telefonnetz bekannten Adressschema. Aus dem breiten Spektrum der Einsatzmöglichkeiten bietet sich insbesondere die Verwendung für Voice over IP Services an.

DynDNS

Es kann nur Rechnern mit fester, sich also nur sehr selten ändernden IP-Adresse ein fester Rechnername zugeordnet werden. Da jedoch sehr viele Nutzer mit Heimrechnern eine variable IP-Adresse haben (mit jeder Einwahl in das Internet wird eine andere IP-Adresse aus einem Pool zugeteilt), gibt es inzwischen DynDNS-Betreiber, die dafür sorgen, dass man auch mit solch rasch ändernden Adressen möglichst immer über den selben Rechnernamen erreichbar ist.

DNS-Sicherheit

Das DNS ist ein zentraler Bestandteil einer vernetzten IT-Infrastruktur. Eine Störung kann erhebliche Kosten nach sich ziehen und eine Verfälschung von DNS-Daten Ausgangspunkt von Angriffen sein. Mehr als zehn Jahre nach der ursprünglichen Spezifikation wurde DNS um Security-Funktionen ergänzt. Folgende Verfahren sind verfügbar:

- Bei TSIG (Transaction Signatures) handelt es sich um ein einfaches, auf symmetrischen Schlüsseln beruhendes Verfahren, mit dem der Datenverkehr zwischen DNS-Servern gesichert werden kann.
- Bei DNSSEC (DNS Security) handelt es sich um ein komplexes Public Key Verfahren, mit dem nahezu alle DNS-Sicherheitsanforderungen erfüllt werden können. Neben der Server-Server-Kommunikation wird auch die Client-Server-Kommunikation gesichert.

Domain-Registrierung

Um DNS-Namen im Internet bekanntmachen zu können, muss der Besitzer die Domain, die diese Namen enthält, registrieren. Durch eine Registrierung wird sichergestellt, dass bestimmte formale Regeln eingehalten werden und dass Domain-Namen weltweit eindeutig sind. Domain-Registrierungen werden von Organisationen (Registrars) vorgenommen, die dazu von der IANA bzw. ICANN autorisiert wurden. Registrierungen sind gebührenpflichtig.

Domain

Eine **Domäne** (englisch *domain*) ist im technischen Zusammenhang die Bezeichnung für eine Gruppe vernetzter Computer oder einen einzelnen Rechner aus einer solchen Gruppe.

Es existieren verschiedene Systeme zur Benennung von Domänen. Das bekannteste ist das Domain Name System (DNS). In diesem nennt man die Namen höchster Ebene Top-Level-Domains. Dementsprechend heißen die Domänen zweiter und dritter Ebene *Second-* bzw. *Third-Level-Domains*.

Eine Domäne ist eine Internetadresse, wie www.example.org. Technisch gesehen ist eine Domäne eine Art Textvariante einer IP-Adresse. Eine IP-Adresse ist eine Nummer, unter der Sie im WWW (World Wide Web) erreichbar sind. Wenn Sie eine Internetseite erreichen, tun Sie dies entweder über eine Domäne oder eine IP-Adresse. Zweck einer Domäne ist es, dass Sie und die Besucher Ihrer Seite sich die Adresse der Seite leichter merken können. Gegenüber der IP z. B. ist www.example.org leichter zu merken und kürzer als 216.23.234.123

Top Level Domain

Jeder Name einer Domain im Internet besteht aus einer Folge von durch Punkten getrennte Namen. Die englische Bezeichnung **Top Level Domain** (übersetzt *Bereich oberster Ebene*; Abkürzung **TLD**) bezeichnet dabei den letzten Namen dieser Folge und stellt die höchste Ebene der Namensauflösung dar. Heißt der Rechner beispielsweise *www.wikipedia.org*, so ist *org* die Top Level Domain dieses Rechnernamens.

Im so genannten Domain Name System (DNS) werden die kompletten Namen und damit auch die TLDs referenziert und aufgelöst, beispielsweise einer eindeutigen IP-Adresse zugeordnet.

TLDs können dabei in zwei Hauptgruppen aufgeteilt werden; allgemeine TLDs (*generic TLD*s; *gTLD*s) und länderspezifische TLDs (*country-code TLD*s; *ccTLD*s). Länderspezifische TLD-Bezeichner bestehen dabei immer aus zwei Buchstaben, allgemeine TLDs setzen sich, mit Ausnahme der TLD .eu, aus drei oder mehr Buchstaben zusammen.

Allgemeine Top Level Domains

.**aero** - für in der Luftfahrt tätige Organisationen – weltweit
.**arpa** - TLD des ursprünglichen Arpanets, jetzt verwendet als Address and Routing Parameter Area
.**biz** - business, nur für Handelsfirmen - weltweit
.**com** - commercial, ursprünglich nur für US-Firmen, jetzt frei für jeden - weltweit (*)
.**coop** - cooperatives - weltweit
.**edu** - educational, nur für Bildungseinrichtungen (*)
.**gov** - government, nur Regierungsorgane der USA (*)
.**info** - Informationsanbieter weltweit
.**int** - international, internationale Regierungsorganisationen (z. B. www.nato.int)
.**mil** - military, nur für militärische Einrichtungen der USA (*)
.**museum** - Museen - weltweit
.**name** - nur für natürliche Personen oder Familien (Privatpersonen) - weltweit
.**net** - Netzverwaltungseinrichtung - weltweit (*)
.**org** - organization, für nichtkommerzielle Organisationen - weltweit (*)
.**pro** - professions (Anwälte, Steuerberater, Ärzte usw.) - nur für gennante Berufsgruppen der USA

(*) = ursprünglich die einzigen TLD, die existierten.

Aufgrund der liberalen Vergaberegeln für die TLD *.com*, *.net*, *.org*, *.info* sowie (mit kleineren Einschränkungen) *.biz* und (neuerdings) *.name* ist die ursprüngliche Bedeutung dieser TLD jedoch weitestgehend abhanden gekommen. Eine solche TLD weist nicht notwendigerweise auf einen entsprechenden Gebrauch hin. Selbiges gilt übrigens für viele ccTLD.

Die *.arpa* TLD sollte ursprünglich nur eine temporäre Lösung bei der Einrichtung des DNS im Internet sein, jedoch stellte sich die spätere Auflösung dieser Domain als unpraktisch heraus. Die Subdomain *in-addr.arpa* ist weltweit im Einsatz um das Auflösen einer IP-Adresse in einen Domainnamen (reverse lookup) zu ermöglichen, eine weitere Subdomain *e164.arpa* wird für ENUM, die Adressierung von Internet-Diensten über Telefonnummern, verwendet.

Länderspezifische Top Level Domains

Jedes Land besitzt genau einen Zwei-Buchstaben Code nach ISO 3166 (Ausnahme: das Vereinigte Königreich besitzt die TLD .uk **und** .gb).

A-B

.ac - Ascension
.ad - Andorra
.ae - Vereinigte Arabische Emirate
.af - Afghanistan
.ag - Antigua und Barbuda
.ai - Anguilla
.al - Albanien
.am - Armenien
.an - Niederländische Antillen
.ao - Angola
.aq - Antarktis (hier definiert als alles südlich von 60°)
.ar - Argentinien
.as - Amerikanisch-Samoa
.at - Österreich
.au - Australien
.aw - Aruba
.az - Aserbaidschan
.ba - Bosnien-Herzegowina
.bb - Barbados
.bd - Bangladesch
.be - Belgien
.bf - Burkina Faso
.bg - Bulgarien
.bh - Bahrain
.bi - Burundi
.bj - Benin
.bm - Bermuda
.bn - Brunei
.bo - Bolivien
.br - Brasilien
.bs - Bahamas
.bt - Bhutan
.bv - Bouvetinsel
.bw - Botswana
.by - Weißrussland (Belorussland)
.bz - Belize

C

.ca - Kanada
.cc - Kokosinseln (auch Keelinginseln)
.cd - Demokratische Republik Kongo
.cf - Zentralafrikanische Republik
.cg - Republik Kongo
.ch - Schweiz
.ci - Côte d'Ivoire
.ck - Cookinseln
.cl - Chile
.cm - Kamerun
.cn - Volksrepublik China
.co - Kolumbien
.cr - Costa Rica
.cs - Serbien und Montenegro ("Srbija i Crna Gora"); wurde früher für die Tschechoslowakei verwendet
.cu - Kuba
.cv - Kap Verde
.cx - Weihnachtsinsel
.cy - Zypern
.cz - Tschechien

D-F

.dd - DDR (wurde definiert, aber nie genutzt, heute nicht mehr existent)
.de - Deutschland
.dj - Djibouti
.dk - Dänemark

.dm - Dominica
.do - Dominikanische Republik
.dz - Algerien
.ec - Ecuador
.ee - Estland
.eg - Ägypten
.eh - Westsahara
.er - Eritrea
.es - Spanien
.et - Äthiopien
.fi - Finnland
.fj - Fidschi
.fk - Falklandinseln (Malvinen)
.fm - Mikronesien
.fo - Färöer
.fr - Frankreich

G-H

.ga - Gabun
.gb - Vereinigtes Königreich
.gd - Grenada
.ge - Georgien
.gf - Französisch-Guayana
.gg - Guernsey
.gh - Ghana
.gi - Gibraltar
.gl - Grönland
.gm - Gambia
.gn - Guinea
.gp - Guadeloupe
.gq - Äquatorialguinea
.gr - Griechenland
.gs - Südgeorgien und die Südlichen Sandwichinseln
.gt - Guatemala
.gu - Guam
.gw - Guinea-Bissau
.gy - Guyana
.hk - Hongkong
.hm - Heard und McDonaldinseln
.hn - Honduras
.hr - Kroatien (Hrvatska)
.ht - Haiti
.hu - Ungarn

I-K

.id - Indonesien
.ie - Irland
.il - Israel
.im - Isle of Man
.in - Indien
.io - Britisches Territorium im Indischen Ozean
.iq - Irak
.ir - Iran
.is - Island
.it - Italien
.je - Jersey
.jm - Jamaika
.jo - Jordanien
.jp - Japan
.ke - Kenia
.kg - Kirgisien
.kh - Kambodscha
.ki - Kiribati
.km - Komoren
.kn - St. Kitts und Nevis
.kp - Nordkorea (Volksrepublik)
.kr - Südkorea (Republik)
.kw - Kuwait
.ky - Kaimaninseln
.kz - Kasachstan

L-M

.la - Laos
.lb - Libanon
.lc - St. Lucia
.li - Liechtenstein
.lk - Sri Lanka
.lr - Liberia
.ls - Lesotho
.lt - Litauen
.lu - Luxemburg
.lv - Lettland
.ly - Libyen
.ma - Marokko
.mc - Monaco
.md - Moldawien
.mg - Madagaskar
.mh - Marshallinseln
.mk - Mazedonien
.ml - Mali
.mm - Myanmar (ehemals Birma/Burma)
.mn - Mongolei
.mo - Macao
.mp - Nördliche Marianen
.mq - Martinique
.mr - Mauretanien
.ms - Montserrat
.mt - Malta
.mu - Mauritius
.mv - Malediven
.mw - Malawi
.mx - Mexiko
.my - Malaysia
.mz - Mozambique

N-P

.na - Namibia
.nc - Neukaledonien
.ne - Niger
.nf - Norfolkinsel
.ng - Nigeria
.ni - Nicaragua
.nl - Niederlande
.no - Norwegen
.np - Nepal
.nr - Nauru
.nu - Niue
.nz - Neuseeland
.nt - reserviert für Neutrale Zonen
.om - Oman
.pa - Panama
.pe - Peru
.pf - Französisch-Polynesien
.pg - Papua-Neuguinea
.ph - Philippinen
.pk - Pakistan

.pl - Polen
.pm - St. Pierre und Miquelon .pn - Pitcairninseln
.pr - Puerto Rico
.ps - Palästinensische Autonomiegebiete (reserviert für Palästina)
.pt - Portugal
.pw - Palau
.py - Paraguay

Q-S

.qa - Katar
.re - Réunion
.ro - Rumänien
.ru - Russland
.rw - Ruanda
.sa - Saudi-Arabien
.sb - Salomonen
.sc - Seychellen
.sd - Sudan
.se - Schweden
.sg - Singapur
.sh - St. Helena
.si - Slowenien
.sj - Svalbard und Jan Mayen
.sk - Slowakei
.sl - Sierra Leone
.sm - San Marino
.sn - Senegal
.so - Somalia
.sr - Suriname
.st - São Tomé und Príncipe
.su - die ehemalige Sowjetunion (wieder verfügbar)
.sv - El Salvador
.sy - Syrien
.sz - Swasiland

T-U

.tc - Turks- und Caicosinseln
.td - Tschad
.tf - Französische Süd- und Antarktisgebiete
.tg - Togo
.th - Thailand
.tj - Tadschikistan
.tk - Tokelau
.tl - Timor-Leste
.tm - Turkmenistan
.tn - Tunesien
.to - Tonga
.tp - Ost-Timor
.tr - Türkei
.tt - Trinidad und Tobago
.tv - Tuvalu
.tw - Taiwan
.tz - Tansania
.ua - Ukraine
.ug - Uganda
.uk - Vereinigtes Königreich
.um - Amerikanisch-Ozeanien
.us - USA
.uy - Uruguay
.uz - Usbekistan

V-Z

.va - Vatikanstadt
.vc - St. Vincent und die Grenadinen (GB)
.ve - Venezuela
.vg - Britische Jungferninseln (GB)
.vi - Amerikanische Jungferninseln (USA)
.vn - Vietnam
.vu - Vanuatu
.wf - Wallis und Futuna (Frankreich)
.wg - reserviert für Westjordanland und Gazastreifen (Palästina)
.ws - Samoa
.ye - Jemen
.yt - Mayotte (Frankreich)
.yu - Ehem. Jugoslawien, Serbien und Montenegro (neu: .cs)
.za - Südafrika
.zm - Sambia
.zw - Simbabwe

.EU TLD

.eu - Europäische Union

Obwohl EU in ISO 3166 für die Europäische Union reserviert ist, handelt es sich hierbei nicht um eine *ccTLD* im eigentlichen Sinn, da die EU kein Land ist. Es ist aber auch keine *gTLD*, da sie nur für ein abgegrenztes Gebiet gilt.

Zweckentfremdungen

Die folgenden Top Level Domains, bei denen es sich eigentlich um ccTLDs handelt, werden global vermarktet.

.**ag** – Aktiengesellschaft (eigentlich Antigua und Barbuda)
.**am** – Radio (Amplitudenmodulation, eigentlich Armenien)
.**bz** – Alternative zu .biz (eigentlich Belize)
.**cc** – (eigentlich Kokosinseln)
.**cd** – Compact Disc (eigentlich Demokratische Republik Kongo)
.**fm** – Radio (Frequenzmodulation, eigentlich Mikronesien)
.**ms** – Message, könnte auch Microsoft suggerieren (eigentlich Montserrat)
.**nu** – (eigentlich Niue)
.**to** – engl. zu, nach (eigentlich Tonga)
.**tk** – (eigentlich Tokelau)
.**tm** – engl. trademark (eigentlich Turkmenistan)
.**tv** – Television (eigentlich Tuvalu)
.**vu** – (eigentlich Vanuatu)
.**ws** – Web-Site (eigentlich Samoa)

Folgende Domains werden (auch) für eine andere Region verwendet, als die, für die sie ursprünglich gedacht waren:

.**by** – Bayern www.bayern.by (eigentlich Weißrussland)
.**sh** – Schleswig-Holstein www.wasser.sh (eigentlich St. Helena)
.**la** – Los Angeles, USA (eigentlich Laos)
.**ca** – California, USA (eigentlich Kanada), auch Catalunya
.**cc** – (ehemalige) Sowjetunion (russ. „Советский Союз“)

Manche Provider bieten auch (in der Regel kostenlos) Subdomains an unter Domains, die zwar keine echten Top-Level-Domains darstellen, aber wie solche aussehen:

.**de.ki** – (NIC.de.ki) – Subdomain *de* unter TLD *ki* (Kiribati)
.**eu.tf** – (UNONIC) – Subdomain *eu* unter TLD *tf* (Frz. Südterritorien)
.**de.vu** – (deNicvu) – Subdomain *de* unter TLD *vu* (Vanuatu)
.**de.ms** – (cydots) – Subdomain *de* unter TLD *ms*

Der Vollständigkeit halber seien hier auch die häufig vorkommenden Subdomains, die von .com abgeleitet werden, genannt:

.**co.at** für Kommerzielle aus Österreich
.**co.uk** für Kommerzielle aus Großbritannien
.**com.br** für Kommerzielle aus Brasilien

Alternative Root-DNS

Es gibt im Internet auch Organisationen, die alternative Namensserver betreiben, über welche zusätzlich zu den oben aufgeführten, quasi-offiziellen, vom ICANN kontrollierten TLDs weitere TLDs verfügbar sind. Ein entscheidender Nachteil dabei ist, daß solche Adressen für herkömmliche Internet-Nutzer nicht erreichbar sind. Auch werden sie von Suchmaschinen wie Google ignoriert. Ein weiterer Nachteil ist, daß die Namensräume zweier Betreiber kollidieren können, wie z. B. bei den .biz-Domains des Pacific Root.

Das Projekt OpenNIC versucht dabei die alternativen Systeme zusammenzuführen, betrachtet jedoch die ICANN-TLDs als vorrangig und akzeptiert weder konfliktende noch private Namensräume.

- OpenNIC-eigene TLDs sind: .glue, .indy, .geek, .null, .oss und .parody
- AlterNIC-TLDs sind: .exp, .llc, .lnx, .ltd, .med, .nic, .noc, .porn und .xxx
- Das Free Community Network verwendet die TLD .fcn
- Pacific Root TLDs, die über OpenNIC-Namensserver erreichbar sind, sind: .ais, .bali, .belize, .bio, .cal, .career, .chem, .children, .costarica, .ind, .job, .lib, .medic, .nomad, .npo, .ppp, .sat, .satcom, .satnet, .scuba, .socal, .stream, .work und .www.

Weblinks

- Nicht-länderspezifische Top Level Domains; ICANN (http://www.icann.org/tlds/)
- http://www.denic.de
- http://www.internic.net
- http://www.domains.de
- http://www.icann.org

Subdomain

Als **Subdomain** (auch *Subdomäne*) bezeichnet man beim Domain Name System eine Domain, welche in der Hierarchie unterhalb einer anderen liegt.

Im allgemeinen Sprachgebrauch sind damit Domains in der dritten oder einer weiteren Ebene gemeint. Domains, die direkt unterhalb der Top Level Domain liegen (z. B. *wikipedia* in *wikipedia.org*), bezeichnet man für gewöhnlich nicht als Subdomains.

So ist etwa in *de.wikipedia.org* der Teil *de* eine Subdomain unter *wikipedia.org*.

Einige Dienstleister bieten Subdomains als Weiterleitung zu einer längeren URI an. Bekannte Beispiele sind Subdomains in der Form *subdomain.de.ki* oder *subdomain.de.vu*.

Root-Server

Root-Server nehmen im Internet Domain Name System-Anfragen von Rechnern aus aller Welt entgegen und leiten zu den autoritativen DNS-Servern der gewünschten Top Level Domain weiter.

Root-Server stehen hierarchisch gesehen an oberster Stelle im **Domain Name System**. Sie werden von verschiedenen Institutionen betrieben und von der **ICANN** koordiniert.

Es gibt 13 dieser Server im Internet ('A' bis 'M') und wird wegen einer Beschränkung im DNS-Protokoll auch nicht mehr geben. Einige bestehen jedoch nicht aus einem, sondern mehreren Computern, die zu einem logischen Server zusammengeschlossen sind. Diese

Computer (*Nodes*) befinden sich an verschiedenen Standorten um die ganze Welt und sind per Anycast über dieselbe IP-Adresse erreichbar.

Derzeit (Stand: Januar 2004) nutzen fünf Root-Server die **Anycast**-Technik. Der vom ISC betriebene Root-Server 'F' besteht aus 19 Computern, die auf fünf Kontinente verteilt sind. 'K', vom RIPE betrieben, besteht aus Computern in London, Amsterdam und seit kurzem auch Frankfurt am Main.

Bevor Anycast eingesetzt wurde, befand sich der überwiegende Großteil der Root-Server aus historischen Gründen in den USA. Dies wurde kritisiert, da es das Domain Name System, welches einen wesentlichen Bestandteil des Internets darstellte, für physische Angriffe verwundbarer machte.

Heute erachten es manche für problematisch, dass alle Root-Server ihre Datenbestände von dem von VeriSign betriebenen 'A'-Root-Server synchronisieren und dass das US-Handelsministerium Einfluss auf den Betrieb desselben hat. Änderungen in der Root-Zone bedürfen eines langwierigen Genehmigungsverfahrens durch dieses.

Mehr Informationen finden sich unter http://root-servers.org

Verbindung zum Internet

Akustikkoppler

Der **Akustikkoppler** ist der Vorgänger des Modem, mit dem Daten über eine Telefonleitung verschickt werden. Im Gegensatz zu einem Modem hat der Akustikkoppler keinen eigenen Anschluss an das Telefonnetz, sondern benötigt ein Telefon dazu. An Mikrofon und Lautsprecher des Telefonhörers werden die entsprechenden Gegenstücke des Akustikkopplers befestigt. Die Anwahl erfolgt über die Wahleinrichtung (Wahlscheibe, Tasten) des Telefons.

Aufgrund der Bauart sind Akustikkoppler extrem störanfällig gegenüber externen Geräuschen. Die Übertragungsraten von Akustikkopplern reichen deshalb nur von 300 bis zu 2.400 Baud. Ein weiterer Nachteil ist, dass Akustikkoppler nicht selbstständig Anrufe entgegen nehmen oder die Leitung trennen können. Aus diesem Grund sind sie nicht für den Betrieb von Mailboxen geeignet.

Modem

Ein **Modem** (zusammengesetztes Wort aus **Mo**dulator/**Dem**odulator) dient dazu, digitale Daten in analoge Signale umzuwandeln und umgekehrt.

Modulationsverfahren wurden seit langem in der Rundfunktechnik und später in der Trägerfrequenztechnik der ehemaligen Deutschen Bundespost eingesetzt (wireless modulation). Inzwischen werden sie auch in sehr stark in der leitergebundenen Kommunikation verwendet (wireline modulation).

Mit einem Modem können digitale Daten über ein Kommunikationsnetz (Telefonnetz, Funknetz) an ein anderes Modem übertragen werden. Dort werden im Gegenzug aus den ankommenden analogen Signalen die digitalen Daten regeneriert.

Neben den üblichen Modem-zu-Modem Verbindungen gibt es auch Modems, die die Möglichkeit bieten, mit Faxgeräten bzw. Sprachgegenstellen zu kommunizieren.

Das Modem ist ein leistungsfähigeres Gerät als der Akustikkoppler. Es ist ein eigenständiges Gerät und braucht kein Telefon. Dadurch entstehen keine Übertragungsfehler durch Nebengeräusche (früher wurden Akustikkoppler deshalb in Kissen eingehüllt).

Im analogen Telefonnetz ist die maximale Übertragungsgeschwindigkeit zwischen zwei Modems aufgrund des Abtasttheorems auf 56 kbit/s begrenzt (V.90). Dies gilt allerdings nur für den Fall, dass das öffentliche Telefonnetz benutzt wird, das die Bandbreite der Verbindung auf 3,1 KHz beschränkt. Auf privaten Leitungen können auch wesentlich höhere Geschwindigkeiten erzielt werden. Moderne Modems verwenden Datenkompression, sodass dadurch auch im Telefonnetz ein höherer Durchsatz als 56 kbit/s möglich ist.

Um schnellere Verbindungen herzustellen, bietet es sich an, ISDN zu nutzen. Noch größere Übertragungsraten auf der TeilnehmerAnschlussleitung bietet xDSL.

ISDN

ISDN (Integrated Services Digital Network) bezeichnet einen Standard für ein leitungsvermitteltes digitales Telekommunikationsnetz, das hauptsächlich zur Übertragung von Telefongesprächen genutzt wird. In Europa ist ISDN die Basis aller leitungsvermittelten Telefonnetze. Auch das GSM-Mobilfunknetz basiert auf ISDN.

Die englische Bezeichnung läßt sich sinngemäß als diensteverbundendes digitales Netz übersetzen. Das bedeutet, dass über nur ein Netz nicht nur der Sprachtelefoniedienst, sondern auch Video- und Datendienste (Teletex, Datex, Telefax, Temex, ...) abgewickelt werden können.

Im Zuge des Internet-Booms wurde ISDN auch zunehmend für die Datenübertragung genutzt, da es verglichen mit der analogen Datenübertragung per Modem schneller und somit auch kostengünstiger ist. Mittlerweile wird dafür aufgrund höherer Bandbreite und geringerer Kosten zunehmend DSL eingesetzt.

ISDN ist in Deutschland flächendeckend verfügbar, dort befinden sich auch etwa ein Fünftel der weltweiten ISDN-Anschlüsse. In den USA ist ISDN nahezu unbekannt, weil die dortigen Telefongesellschaften andere digitale Vermittlungstechniken einsetzen.

Unterschiede zum analogen Anschluss

Im Unterschied zu einem analogen TelefonAnschluss stehen bei einem ISDN-BasisAnschluss zwei Leitungen zur Verfügung, die beliebig für Telefongespräche, Fax und Datenübertragungen genutzt werden können. Außerdem können für einen Anschluss bis zu 10 Telefonnummern (genannt MSN) vergeben werden, die beliebig auf die ISDN-Endgeräte verteilt werden können. Durch Dienstkennungen kann eine Nummer jedoch auch für verschiedene Anwendungen genutzt werden, z.B. Telefonie und ISDN-Datenübertragung.

Um analoge Endgeräte (Telefon, Fax, Anrufbeantworter oder Modem) an einen ISDN-Anschluss anzuschließen, benötigt man einen a/b-Wandler oder eine ISDN-Telefonanlage.

Öffentlich verfügbare Anschlusstypen

Ein ISDN-Anschluss ist in zwei Varianten verfügbar: Als Basisanschluss oder als Primärmultiplexanschluss.

Basisanschluss

Der Basisanschluss hat 2 Nutzkanäle. Ein Nutzkanal (auch B-Kanal genannt) bietet eine Datenrate von 64 kbit/s.

Die Basisanschlüsse sind verfügbar als

- Mehrgeräteanschluss (Point-to-Multipoint) zum Anschluss von bis zu 8 ISDN-Endgeräten
- Anlagenanschluss (Point-to-Point) zum Anschluss einer einzigen Telekommunikationseinrichtung, wie z. B. einer TK-Anlage

Primärmultiplexanschluss

Der Primärmultiplexanschluss hat 30 Nutzkanäle. Er ist nur als Anlagenanschluss verfügbar und wird hauptsächlich von größeren Firmen zum Anschluss von Telefonanlagen genutzt.

Implementierungen

In Deutschland wurde ursprünglich ISDN nach dem Standard 1TR6 angeboten, seit 1991 existiert jedoch ein europaweit einheitlicher ISDN-Standard (E-DSS-1). In USA und Japan existiert eine andere Implementierung, die sich im wesentlichen durch die niedrigere Datenrate (56 kbit/s) unterscheidet.

Leistungsmerkmale

Die Telekom bietet den ISDN-Mehrgeräteanschluss in zwei Varianten an, die sich im Bezug auf die angebotenen Leistungsmerkmale unterscheiden: Als Standardanschluss oder als Komfortanschluss.

beim Standardanschluss

- Übermittlung der Rufnummer
- Rückruf bei Besetzt (CCBS)
- Anklopfen (CW)
- Rückfragen/Makeln
- Dreierkonferenz
- Rufnummernanzeige (CLIP)
- Umstecken am Bus (TP): Ein Gespräch kann in der Vermittlungsstelle geparkt werden und das ISDN-Telefongerät an eine andere ISDN-Dose innerhalb des Anschlusses umgesteckt werden.
- bei Anlagenanschluss: Durchwahl (DDI): Es kann ein Nummernblock definiert werden, der direkt angewählt werden kann.

- bei Mehrgeräteanschluss: bis zu 10 Mehrfachrufnummern MSN. Ein ISDN-Anschluss kann also unter bis zu 10 Telefonnummern erreichbar sein. Diese Nummern können flexibel auf die Endgeräte aufgeteilt werden.

ZUSÄTZLICHE MERKMALE BEIM KOMFORTANSCHLUSS

- Anzeige der Rufnummer des Angerufenen (COLP)
- Rückruf bei Nichtmelden (CCNR)
- Anrufweiterschaltungsvarianten
- T-Net-Box
- Tarifinformationen

ÜBERTRAGUNGSVERFAHREN

SPRACHÜBERTRAGUNG

Sprachdaten werden für die Übertragung per Euro-ISDN mit einer Abtastrate von 8 kHz digitalisiert (**P**ulse **C**ode **M**odulation PCM) und mit einer logarithmischen Kennlinie (ITU-T-Standard G.711, A-law) von 12 auf 8 Bit pro Abtastwert komprimiert, um die Besonderheiten der menschlichen Wahrnehmung zu berücksichtigen. Übertragen wird der Frequenzbereich von 300 bis 3400 Hz.

DATENÜBERTRAGUNG

Zur Datenübertragung werden verschiedene Datenübertragungsprotokolle eingesetzt (z. B. X.75, V.110).

SIGNALISIERUNG

Die Signalisierung funktioniert bei ISDN *Out-of-Band* - sie wird also auf einem eigenen Kanal übertragen und nicht wie beim Mehrfrequenzwahlverfahren im Sprachkanal. Dadurch funktioniert der Verbindungsaufbau störungsfreier und schneller.

Technisch wird für die Signalisierung der D-Kanal genutzt, der bei Basisanschlüssen eine Bandbreite von 16 kbit/s und bei Primärmultiplexanschlüssen von 64 kbit/s hat.

SCHNITTSTELLEN

Ein ISDN-Anschluss besteht aus zwei Teilen: Der Teilnehmeranschlussleitung (beim Basisanschluss die U_{K0}-Schnittstelle; beim Primärmultiplexanschluss die U_{K2}-Schnittstelle) und der In-House-Verkabelung (beim Basisanschluss der S_0-Bus; beim Primärmultiplexanschluss die S_{2M}-Schnittstelle). Die Teilnehmeranschlussleitung wird durch einen Netzab-

schluss abgeschlossen (beim Basisanschluss NTBA; beim Primärmultiplexanschluss NTPM).

Die Schnittstelle zu Computersoftware wird meistens durch die CAPI hergestellt.

Digital Subscriber Line (DSL)

Digital Subscriber Line (**DSL**, englisch: „Digitale Teilnehmeranschlussleitung“) bezeichnet verschiedene Techniken für eine vergleichsweise breitbandige digitale Verbindung über ein Zugangsnetz.

Die grundlegende Idee der DSL-Techniken besteht darin, dass die Übertragungskapazität der Kupferdoppeladern des Telefonnetzes bzw. ISDN-Netzes mit der Sprachübertragung nur zu einem geringen Bruchteil ausgenutzt ist. Die für Sprache benötigten Frequenzen betragen im Maximum 130 KHz, über die Doppelader lassen sich aber problemlos auch Frequenzen von 1 MHz und höher übertragen. Diese höheren Frequenzen können mit fortgeschrittenen Leitungscodes und Modulationsverfahren benutzt werden und so zusätzliche Bandbreite für digitale Datenübertragung zur Verfügung stellen.

In der Folge wurden eine Reihe solcher Verfahren der Übertragungstechnik entwickelt. In Deutschland wurde die Bezeichnung DSL jedoch als Synonym für einen breitbandigen Internetzugang (meist über ADSL) bekannt, so dass auch in der Folge andere breitbandige Internetzugänge (z. B. über Satellit) als „DSL“ vermarktet werden. Die DSL-Techniken wurden jedoch auch für andere Anwendungen als den Internetzugang konzipiert.

Nach einer Bitkom-Studie vom Anfang Februar 2003 gibt es 3,2 Millionen DSL-Anschlüsse in Deutschland.

Arten des DSL-Verfahren

Es gibt verschiedene Arten von DSL-Techniken die unter der Bezeichnung „xDSL“ zusammengefasst werden:

- ADSL - Asymmetric Digital Subscriber Line, eine asymmetrische Datenübertragungstechnologie mit Bitraten bis 8 Mbit/s zum Teilnehmer (downstream) und 1 Mbit/s in der Gegenrichtung (upstream);
- HDSL - High Data Rate Digital Subscriber Line, eine asymmetrische Datenübertragungstechnologie mit Datenraten zwischen 1,54 und 2,04 Mbit/s;
- SDSL - Symmetrical Digital Subscriber Line, eine symmetrische Datenübertragungstechnologie mit Bitraten von bis zu 2,3 Mbit/s symmetrisch;
- VDSL - Very High Speed Digital Subscriber Line, eine asymmetrische Datenübertragungstechnologie mit Bitraten von 12,9 bis 51,8 Mbit/s (downstream) bzw. 1,6 bis 2,3 Mbit/s (upstream);

- RADSL - Rate Adaptive Digital Subscriber Line eine asymmetrische Datenübertragungstechnologie mit Bitraten von 6 MBit/s (downstream) bzw. 640 kbit/s (upstream).

Beim ISDN handelt es sich nicht um ein DSL-Verfahren, sondern um ein ein Kommunikationsnetz, das Wählverbindungen ermöglicht. Man könnte allenfalls die Verbindung zwischen NTBA und Vermittlungsstelle (U_{K0} beim Basisanschluss) als DSL-Verfahren bezeichnen.

Andere als DSL bezeichnete Verfahren

- IDSL - *ISDN Digital Subscriber Line* verwendet vorhandene ISDN-Technik und ermöglicht Datenraten bis zu 160 kBit/s.
- skyDSL - Markenname der Strato AG für einen Internetzugang über Satellit.
- T-DSL via Satellit - Markenname der T-Com für einen Internetzugang über Satellit.

Reichweite

xDSL ist aufgrund der physikalischen Eigenschaften der Leitung in der Reichweite begrenzt. Generell gilt: Je höher die Bitrate, um so geringer die Reichweite. Für alle xDSL-Varianten sind daher Modi definiert, mit denen durch Verringerung der Bitrate - teilweise sogar dynamisch adaptiv - die Reichweite erhöht werden kann.

Anwendungen

Während ISDN in erster Linie für die Telefonie mit zwei Amtsleitungen genutzt wird, ist ADSL die erste Technologie, die Netzbetreiber für den schnellen Internet-Zugang von Privatkunden installiert haben. SDSL ist für beide Bereiche geeignet und kommt hauptsächlich für Geschäftskunden zum Einsatz.

Die Tendenz geht dahin, mehrere Dienste über eine einzige Doppelader übertragen zu können - idealerweise das "Triple Play" aus Telefonie, Internet-Zugang und Video.

DSL-Geräte

Für den xDSL-Zugang werden (sowohl auf Kunden- als auch auf Seite der Telefongesellschaft) folgende Hardwarebauteile benötigt:

- DSL-Modem oder ATU-R (**A**DSL **T**ransceiver **U**nit - **R**emote)
- Splitter oder BBAE
- DSLAM (**D**igital **S**ubscriber **L**ine **A**ccess **M**ultiplexer) oder ATU-C (**A**DSL **T**ransceiver **U**nit - **C**entral Office)
- DSL-AC (**D**igital **S**ubscriber **L**ine **A**ccess **C**oncentrator) oder auch Breitband-PoP

Mit **ATU-C** wird das Modem in der Vermittlungsstelle bezeichnet, wobei das C für Central site steht

Dazu kann je nach technischer Realisierung weiteres Equipment wie RADIUS-Server für die Benutzeranmeldung, -Verwaltung und Billing (Verbrauchsdatenspeicherung zum Zwecke der Rechnungserstellung) oder Splitter zur Abtrennung von ISDN/POTS-Signalen kommen. Im erweiterten Sinne gehört auch noch der PC/Router des Kunden zur DSL-Ausrüstung, weil dort die PPPoE-Strecke vom DSL-AC terminiert.

Protokolle

Protokolle für xDSL-Technologien sind beispielsweise:

- PPPoE (*PPP over Ethernet*) - Protokoll, das die Kapselung von PPP-Paketen in Ethernet-Frames regelt; PPPoE wird unter anderem von der Telekom für T-DSL verwendet.
- PPPoA (*PPP over ATM*) - Protokoll, das die Kapselung von PPP-Paketen in ATM-Zellen regelt.
- PPTP (*Point-to-Point Tunneling Protocoll*) - Protokoll, das einen Tunnel über eine PPP-Verbindung herstellt. (selten für DSL verwendet, z. B. in München von M"net (*http://www.m-net.de/*), in Österreich aber häufiger)

Point-to-Point Protocol

Das **Point-to-Point Protocol** bzw. **Punkt-zu-Punkt-Protokoll** (PPP) ist ein Protokoll zum Verbindungsaufbau über Wählleitungen (zumeist über Modem oder ISDN).

<table>
<tr><td>Anwendung</td><td>FTP</td><td>SMTP</td><td>HTTP</td><td>DNS</td><td>...</td></tr>
<tr><td>Transport</td><td colspan="3">TCP</td><td colspan="2">UDP</td></tr>
<tr><td>Netzwerk</td><td colspan="5">IP</td></tr>
<tr><td rowspan="2">Netzzugang</td><td colspan="5">PPP/PPPoE</td></tr>
<tr><td colspan="3">Serielle Leitung</td><td>Modem</td><td>...</td></tr>
</table>

PPP mit TCP/IP-Protokollstapel

Seltener wird PPP für statische Verbindungen (Standleitungen) verwendet, beispielsweise um die Authentifizierungs-Mechanismen (PAP, CHAP) zu nutzen. Hierfür kommen meist modifizierte Protokolle wie PPPoE oder PPTP zum Einsatz.

PPP ist heute das Standardprotokoll, das Internet-Provider für die Einwahl der Kunden verwenden, die Spezifikationen sind jedoch so definiert, dass PPP nicht ausschließlich TCP/IP-Verbindungen unterstützt.

PPP over Ethernet

PPPoE steht für "PPP over Ethernet", also die Nutzung des Netzwerkprotokolls PPP über eine Ethernet-Verbindung. PPPoE wird heute bei ADSL-Anschlüssen in Deutschland verwendet.

Österreich z. B. verwendet für ADSL-Zugänge das *Point To Point Tunneling Protocol* (PPTP).

Motivation für die Entwicklung von PPPoE war, die Möglichkeiten von PPP wie Authentifizierung und Netzwerkkonfiguration (IP-Adresse, Gateway) auf dem schnelleren Ethernet zur Verfügung zu stellen. Modems werden üblicherweise über eine serielle Schnittstelle mit dem Rechner verbunden, die Datenraten von maximal 192 kBit/s erlaubt. Die Einwahl über PPP gestattet hier aber eine automatische Netzwerkkonfiguration und bei Bedarf auch eine Zugangskontrolle über einen Benutzernamen und Passwort.

Für ADSL (8 MBit/s Downstream) ist eine serielle Verbindung zu langsam, ein ADSL-Modem benötigt also eine schnellere Verbindung. Da Ethernetkabel mindestens 10 MBit/s unterstützen, eignen sie sich gut dafür. Der normale TCP/IP-Protokollstapel bietet jedoch hinsichtlich Zugangskontrolle und Netzwerkkonfiguration wenig Möglichkeiten. PPPoE ergänzt diese Fähigkeiten. Auf dem Rechner muss dazu ein spezieller Treiber für dieses Protokoll installiert sein.

Ein Problem bei der Verwendung von PPPoE ist die verringerte maximale Paketgröße (Maximum Transfer Unit, MTU), die bei Ethernet grundsätzlich 1500 Byte beträgt. Bei PPPoE verringert sie sich jedoch wegen eines zusätzlichen Headers von 8 Byte auf 1492 Byte. Falls der TCP/IP-Treiber die Größe beim Senden nicht ermitteln kann, werden trotzdem 1500 Byte große Datenpakete erzeugt. Dies ist normalerweise kein Problem, da das Internet-Protokoll das Paket fragmentieren kann. Fragmentierung wird wegen des erforderlichen Aufwandes jedoch zunehmend im Internet abgeschaltet, so dass ohne besondere Maßnahmen manche Webserver nicht zugänglich erscheinen.

Standleitung

Eine **Standleitung** ist eine permanente digitale Datenverbindung, die von einem Netzbetreiber zur Verfügung gestellt wird. Im Gegensatz zu einer Wählverbindung steht der gesamte Übertragungsweg immer zur Verfügung. Standleitungen haben keine Rufnummer. In der Regel wird die Standleitung gemietet, ihr Preis ist unter anderem abhängig von der zu überbrückenden Entfernung. Der Anschluss einer Standleitung ist bittransparent. Es können Daten jeder Art übertragen werden.

In der Fachsprache der Computertechnik wird eine Standleitung manchmal auch als **DDV** (**D**aten**d**irekt**v**erbindung) bezeichnet. Die Deutsche Telekom hat den Ausdruck **Direktrufnetz** geprägt, eine Standleitung wurde als ein **Hauptanschluss für Direktruf** (HfD) bezeichnet.

Standleitungen sind heute mit Datenraten von 50 Bit/sec bis zu 2 GBit/sec von allen größeren Netzbetreibern zu mieten. Die Ausfallsicherheit der Standleitungen auf Kupferkabeln ist wesentlich höher als die von ISDN-Anschlüssen. Eine noch höhere Stabilität kann man durch den Einsatz von Glasfaser-Standleitungen erreichen.

Für Standleitungen werden die verschiedensten Verfahren und Schnittstellen eingesetzt. Der Anschluss an eine Standleitung hat häufig eine X.21- oder G.703-Schnittstelle. Auf diesen Schnittstellen setzen dann, an den beiden Endpunkten einer Standleitung, die jeweiligen Router der verbundenen Netzwerke an.

Im weitaus häufigsten Anwendungsfall werden aber nicht Router über eine Standleitung verbunden, sondern die Nebenstellenanlagen einer Firma, deren Gebäude an verschiedenen Standorten einer Stadt stehen, bzw. die Filialen in mehreren Städten hat.

Wichtige Standleitungen werden im Normalfall immer sekundär abgesichert, dies geschieht in der Regel durch alternative Datenwege, also andere Standleitungen, oder durch ISDN-Wählleitungen, auf die im Fehlerfall umgeschaltet werden kann.

Wird über eine Standleitung ein Netzwerk mit dem Internet verbunden, so beinhaltet dies in den meisten Fällen auch die Zuweisung mindestens einer festen öffentlichen IP-Adresse. Diese Standleitungen eignen sie sich dann besonders um Server im Internet zu betreiben. Genauso häufig werden Standleitungen jedoch auch für die Verbindung von zwei privaten (Teil-)Netzen genutzt, hier werden dann keine öffentlichen IP-Adressen zugewiesen. Auch wird bei Standleitungen der zweiten Art nur die Bereitstellung durch den Anbieter berechnet, während bei Standleitungen ins Internet neben der Bereitstellung auch der reale Datentransfer zu bezahlen ist.

Häufig wird das von der Telekom angebotene T-DSL mit einer Standleitung verwechselt. Das ist aber falsch, da hier die Verbindung regelmäßig vom Provider getrennt wird. Außerdem werden die IP-Adressen dynamisch adressiert. Durch die Verwendung dynamischer DNS-Dienste (DynDNS) kann eine quasi-Standleitung mit einem festen Domänennamen eingerichtet werden.

WIRELESS LAN

Wireless LAN (Wireless Local Area Network, kurz WLAN) steht für "drahtloses lokales Netzwerk", wobei meistens der Standard IEEE 802.11 gemeint ist. Das Kürzel "Wi-Fi" wird in diesem Zusammenhang oft genannt. Es steht für Wireless Fidelity. Verschiedene große Hersteller haben sich zusammengeschlossen und testen ihre Geräte auf Interoperabilität. Geräte, die diesen Test bestehen, bekommen das Wi-Fi Siegel.

SPEZIFIKATION

Zum Standard IEEE 802.11 gibt es folgende zusätzliche Standards:

IEEE 802.11 FHSS

- Datenrate: 2 MBit/sec brutto

- Frequency Hopping Spread Spectrum (Wechsel zwischen den Kanälen) während des Sendens
- 13 Kanäle in Europa, 11 in USA
- max. Sendeleistung 20 dbm EIRP
- 2,4 GHz

IEEE 802.11 DSSS

- Datenrate: 2 MBit/sec brutto
- Direct Sequence Spread Spectrum (Datenübertragung immer auf einem Kanal)
- 13 Kanäle in Europa (3 total überlappungsfrei, 4 ebenfalls möglich), 11 in USA (3 überlappungsfrei)
- max. Sendeleistung 20 dbm EIRP
- 2,4 GHz

IEEE 802.11a

- Datenrate: 54 MBit/sec brutto (ca. 24 MBit/s netto)
- 8 überlappungsfreie Kanäle (bis zu 12 möglich)
- max. Sendeleistung 1 Watt
- Signalmodulation: OFDM (Orthogonal Frequency Division Multiplexing), da im Frequenzbereich von über 5 GHz DSSS nicht ausreicht
- 5 GHz
- in Europa nur mit starken Einschränkungen erlaubt

IEEE 802.11b

- Datenrate: 11 MBit/sec brutto (ca. 6 MBit/s netto)
- Direct Sequence Spread Spectrum (Datenübertragung immer auf einem Kanal)
- 13 Kanäle in Europa (3 total überlappungsfrei, 4 ebenfalls möglich), 11 in USA (3 überlappungsfrei)
- max. Sendeleistung 20 dbm EIRP
- Signalmodulation: CCK (Complementary Code Keying)
- 2,4 GHz

IEEE 802.11g

- Datenrate: 54 MBit/sec brutto (ca. 23 MBit/s netto)
- Direct Sequence Spread Spectrum (Datenübertragung immer auf einem Kanal)
- 13 Kanäle in Europa (3 total überlappungsfrei, 4 ebenfalls möglich), 11 in USA (3 überlappungsfrei)
- max. Sendeleistung 20 dbm EIRP
- Signalmodulation: OFDM (Orthogonal Frequency Division Multiplexing)
- Abwärtskompatibel zu 802.11b
- 2,4 GHz

IEEE 802.11h

- Datenrate: 54 MBit/sec brutto (ca. 24 MBit/s netto)
- Direct Sequence Spread Spectrum (Datenübertragung immer auf einem Kanal)
- 8+11 Kanäle (überlappungsfrei)
- max. Sendeleistung 1 Watt
- 5 GHz
- Signalmodulation: OFDM (Orthogonal Frequency Division Multiplexing)
- 802.11h = 802.11a + Transmit Power Control + Dynamic Frequncy Change
- in Europa erlaubt

Sicherheit

Teil des WLAN-Standards IEEE 802.11 ist *Wired Equivalent Privacy* (WEP), ein Sicherheitsstandard, der FHSS, DSSS und den RC4-Algorithmus enthält. Die enthaltene Verschlüsselung mit nur 40 Bit reicht jedoch längst nicht aus. Durch das Sammeln von Schlüsselpaaren sind *Known-Plaintext-Attacken* möglich. Jeder Nutzer des Netzwerkes kann den gesamten Verkehr zudem mitlesen. Die Kombination von RC4 und CRC wird als mathematisch unsicher betrachtet.

Aus diesen Gründen haben sich technische Ergänzungen entwickelt, etwa WEPplus mit 128 Bit- Verschlüsselung, Fast Packet Keying, Extensible Authentication Protocol, Kerberos oder High Security Solution, die alle mehr oder weniger gut das Sicherheitsproblem von WLAN verkleinern. Auch der Einsatz von Zugangsschutzsystemen (*firewalls*) ist gerade bei WLAN unabdinglich.

Der Nachfolger des WEP ist der WPA-Standard. Er bietet eine erhöhte Sicherheit und gilt zur Zeit als nicht zu entschlüsseln, solange man bei der Einrichtung keine trivialen Passwörter verwendet (*Brute Force Attacke*).

Eine alternative Herangehensweise besteht darin die Verschlüsslung komplett auf IP-Ebene zu verlagern. Hierbei wird der Datenverkehr beispielsweise durch einen VPN-Tunnel (EDV) geschützt. Besonders in freien Funknetzen werden so die Inkompatiblitäten verschiedener Hardware umgangen, eine zentrale Benutzerverwaltung vermieden und der offene Charakter des Netzes gewahrt.

Beim so genannten WarWalking oder WarDriving werden mit einem WLAN-fähigen Notebook oder PDA offene WLAN-Netze gesucht. Diese werden dann mit Kreide markiert (WarChalking). Das Ziel ist hierbei entweder, Sicherheitslücken aufzudecken und dem Betreiber zu melden, oder aber einen kostenlosen Internetzugang zu erhalten oder gar Daten auszuspähen oder zu manipulieren. Fährt man bei der Suche eines WLAN-Netzes mit einem Auto, so spricht man von WarDriving.

Topologie

Netze mit Wireless LAN nach IEEE 802.11 arbeiten meistens im Infrastruktur-Modus, bei der eine oder mehrere Basisstationen (Access Points) die Kommunikation zwischen den Clients organisieren. Der Datentransport läuft immer über die Basisstation(en). Seltener werden heute Netze im Ad-Hoc-Modus betrieben, bei dem die Clients direkt miteinander kommunizieren.

Literatur

- Telepolis-Buch: Armin Medosch: Freie Netze. Offene WLAN-Zugangsknoten: Kultur, Politik und Ökonomie. (*http://www.heise.de/tp/deutsch/html/buch_11.html*), ISBN 3936931100

Internetdienstanbieter / Provider

Ein **Internetdienstanbieter** (englisch *Internet Service Provider*, ISP; im deutschsprachigen Raum auch oft nur *Provider* genannt) bietet gegen Entgelt verschiedene technische Leistungen an, die für die Nutzung oder den Betrieb von Internet-Diensten erforderlich sind.

Die Leistungen werden grob in zwei Bereiche unterteilt: *Hosting* und *Zugang*. Nur wenige ISPs decken beide Bereiche komplett ab. Die meisten, insbesondere kleinere Internetdienstanbieter, beschränken sich sogar auf sehr kleine Teilbereiche. Größere Provider bieten hingegen sogar einzelne Produkte an, die Leistungen aus beiden Bereichen umfassen.

Hosting

Die wichtigsten Leistungen aus dem Bereich Hosting sind Registrierung und Betrieb von Domains, Vermietung von Webservern (komplett oder teilweise) und Vermietung von Platz in einem Rechenzentrum inklusive Internet-Anbindung, Strom- und Notstromversorgung etc.

Die bedeutendsten Hosting-Provider für Privatkunden in Deutschland sind Strato und Puretec.

Zugang

Die wichtigsten Leistungen von Zugangsanbietern (auch *Access-Provider* genannt) sind die Bereitstellung von Wählverbindungen, Breitbandzugängen und Standleitungen.

Die bedeutendsten Access-Provider für Privatkunden in Deutschland sind T-Online, AOL und Freenet.

Client-Server-System

Ein **Client-Server-System** oder auch **Klient-Server-System** besteht aus einem Client, der eine Verbindung mit einem zentralen Server aufbaut.

- Der Client bietet die Benutzeroberfläche oder die Benutzerschnittstelle der Anwendung an.
- Der Server stellt die Funktionalität zur Verfügung.

In der Anfangszeit, als Computer noch ganze Räume füllten, war dies der übliche Aufbau einer Anwendung. In jedem Büro stand ein Terminal, welches man als "Thin-Client" bezeichnen könnte, welches eine Benutzeroberfläche für die Applikation anbot, die auf dem Server lief.

Mit der Verbreitung der Personal-Computer wurde immer mehr Rechenkapazität auf die Bürorechner ausgelagert. Der Server in einem solchen Umfeld bietet meistens nur noch die Daten an. Eine häufige Form ist z.B. ein zentraler Datenbankserver. Ein solches Client-Programm wird auch als Fat-Client bezeichnet. Mit der zunehmenden Verbreitung von Intranet geht dieser Trend wieder zurück. Hier ist dann der Browser der Thin-Client, die eigentliche Programmlogik liegt auf einem Application Server.

Network Address Translation

NAT (**Network Address Translation**) ist in Computernetzwerken ein Verfahren, bei dem private IP-Adressen auf öffentliche IP-Adressen abgebildet werden. Werden auch die Port-Nummern umgeschrieben spricht man dabei von **maskieren**.

Verwendung

NAT wird aus verschiedenen Gründen verwendet. Hauptsächlich ist NAT notwendig, weil öffentliche IP-Adressen immer knapper werden und man deshalb private IP-Adressen einsetzen muss. Zum Anderen kann es der Datensicherheit dienen, weil die interne Struktur des Netzwerks nach außen hin verborgen bleibt (Security through Obscurity).

Funktionsweise

Ein NAT-Gerät verbindet mit zwei Netzwerkkarten das lokale Netz mit dem Internet. Man unterscheidet zwischen *Source NAT*, bei dem die Quell-IP-Adresse ersetzt wird, und *Destination NAT*, bei dem die Ziel-IP-Adresse ersetzt wird. Bei **Basic NAT** wird jede interne IP durch eine externe IP ersetzt. Man spricht deshalb von einer 1:1-Übersetzung.

Beispiel: Öffentliche verfügbare Adressen: `205.0.0.0/24`

Source NAT:

Quell-IP	Ziel-IP	**Router**	**Quell-IP**	**Ziel-IP**
192.168.0.2	*170.0.0.1*	**----------> NAT**	205.0.0.2	*170.0.0.1*
192.168.0.3	*170.0.0.1*		205.0.0.3	*170.0.0.1*

Bei ausgehenden Paketen wird die (private) Quell-IP-Adresse durch eine noch nicht benutzte (öffentliche) IP ersetzt. Zusätzlich merkt sich der Router mittels einer Tabelle die Quell- und Ziel-IP-Adresse:

- `192.168.0.2 <-> 170.0.0.1`
- `192.168.0.3 <-> 170.0.0.1`

Destination NAT:

Quell-IP	Ziel-IP		Quell-IP	Ziel-IP
170.0.0.1	*205.0.0.2*	**Router ----------> NAT**	170.0.0.1	*192.168.0.2*
170.0.0.1	*205.0.0.3*		170.0.0.1	*192.168.0.3*

Bei eingehenden Paketen kann anhand der Quell-IP-Adresse und des Tabelleneintrags festgestellt werden, welcher Computer die Pakete angefordert hatte (hier: `192.168.0.2` und `192.168.0.3`). Der Router kann dadurch die (öffentliche) Ziel-IP durch die ursprüngliche Quell-IP `192.168.0.2` bzw. `192.168.0.3` austauschen.

Für den Host im internen Netz (z.B. `192.168.0.2`) sind diese Vorgänge transparent, d.h. er bekommt von der Adressumsetzung nichts mit. Es können auch Verbindungen von extern nach intern aufgebaut werden. Ein Host könnte somit auch als Server dienen.

Masquerading ist eine Implementation von NAPT (Network Address Port Translation), bei dem auch die Ports umgeschrieben werden.

Proxy

Ein **Proxy** oder **Proxyserver** (vom engl. proxy = Stellvertreter) ist ein Programm, das zwischen Server und Client vermittelt. Dem Server gegenüber verhält sich das Programm wie ein Client, dem Client gegenüber wie ein Server.

Funktion

Im einfachsten Fall leitet der Proxy die Daten einfach weiter, üblicherweise hat ein Proxy aber eine der folgenden Funktionen:

- *Zwischenspeicher (Cache)*: Der Proxy speichert häufig gestellte Anfragen und kann diese dann beantworten, ohne zuerst den Server zu fragen. Dadurch können Anfragen schneller beantwortet werden, und es wird gleichzeitig die Netzlast verringert.
- *Filter*: Mittels Proxy können beispielsweise bestimmte Kategorien von Webseiten für den Benutzer gesperrt werden. Es kann auch der Inhalt auf schädliche Programme durchsucht werden. Somit ist ein Proxy auch oft Teil von Firewalls.
- *Ermöglichung des Zugriffs*: Ist der Server nicht frei im Internet erreichbar, so kann ein vorgeschalteter Proxy den Zugriff ermöglichen. Ein Angreifer kann dann den Server nicht mehr direkt angreifen, sondern nur den Proxy. Es kann auch der Zugriff von Clients auf Webserver nur über einen Proxy ermöglicht werden.
- *Vorverarbeitung von Daten*: Proxys können auch gewisse Applikationsfunktionen übernehmen, beispielsweise Daten in ein standardisiertes Format bringen.

- *Anonymisierungsdienst*: Der Proxy leitet die Daten des Clients zum Server weiter, wodurch der Server die IP-Adresse des Clients nicht auslesen kann.

Protokolle

Proxys sind generell für jedes verbindu*ngsorientierte Protokoll möglich. Häufig werden sie für folgende Protokolle verwendet:*

- HTTP: Die meisten Provider bieten Ihren Kunden die Verwendung eines Proxies an. Dadurch wird die Netzlast verringert und der Zugriff beschleunigt. In Firmen hingegen wird über solche Proxies oft das Surfverhalten der Mitarbeiter eingeschränkt bzw. kontrolliert.
- FTP: Die meisten HTTP-Proxies beherrschen auch FTP. Hier sind dieselben Funktionen wie bei HTTP möglich.
- SMTP: Manche Firewalls bieten einen SMTP-Proxy an, der den Mailverkehr zwischen Internet und Mailserver überwacht und bestimmte gefährliche bzw. unerwünschte Befehle ausfiltert.
- *Applikationsproxy: Ein Proxy, der auf ein bestimmtes Server-Progra*mm zugeschnitten ist, und nur dessen Protokoll erkennt. Diese Form eines Proxys wird oft dazu verwendet, den eigentlichen Server in ein geschütztes Netz zu stellen und nur durch den Proxy erreichbar zu machen. Auf diese Art ist der Server weitgehend vor Angriffen geschützt. Die Proxy-Software ist weit weniger komplex, und daher auch sicherer gegen Angriffe.

Sonderformen

- **Transparenter Proxy**: Die Verwendung eines Proxy-Servers muss meist dem Client explizit mitgeteilt werden. Ein transparenter Proxy muss hingegen nicht explizit angegeben werden. Ein Gateway erkennt die Verwendung des vom Proxy verwendeten Protokolls und leitet die Anfragen an den Proxy weiter, ohne dass das Anwendungsprogramm etwas davon bemerkt.
- **Reverse Proxy**: Tritt statt dem eigentlichen Server in Erscheinung. Dadurch können etwa Zugriffskontrollen oder die Verteilung der Last für die Webserver realisiert werden.

Proxy Software

Bekannte Proxy-Server-Software:

- WebWasher (*http://www.webwasher.com/client/home/index.html?lang=de_DE*) - lokaler *HTTP*-Proxy (Windows)
- Proxomitron (*http://www.proxomitron.info/*) ehemals *Junkbuster* (*http://www.junkbusters.com/*) - lokaler HTTP-Proxy (Windows, Unix)

- Squid (*http://www.squid-cache.org/*) (Unix)
- iProxy (*http://www.research.att.com/sw/tools/iproxy*) (AT&T)
- Java Anonymity & Privacy (*http://anon.inf.tu-dresden.de/*) (JAP) - Anonymisierungsprogramm

ROUTING

Beim **Routing** (*amerikanische Aussprache* etwa wie "Rauting" und *britische Aussprache* etwa wie "Ruting"), einem Begriff der Netzwerktechnik, wird dafür gesorgt, dass logisch adressierte Pakete aus dem Ursprungs-Netz heraus kommen und in Richtung ihres Ziel-Netzes weitergeleitet werden. Routing ist die Basis des Internet. Ohne Routing würde das Internet nicht existieren, und alle Netze wären autonom. Die Datenpakete können dabei viele verschiedene Zwischen-Netzwerke auf dem Weg dorthin passieren. Routing passiert auf Layer 3 des OSI-Modells.

Hubs und Switches bewegen Daten nur im lokalen Netzwerk, wohingegen der Router auch Nachbar-Netzwerke kennt. Dieser Artikel beschreibt Routing auf eine Hardware-unabhängige Art. Für Informationen über Router selbst siehe den Router-Artikel.

Um zu wissen, wohin Pakete gesendet werden sollen, muss man die Struktur des Netzwerks kennen. In kleinen Netzwerken kann das Routing sehr einfach sein und wird oft per Hand konfiguriert. Große Netzwerke können eine komplexe Topologie haben, die sich möglicherweise auch noch häufig ändert, was auch das Routing zu einer komplexen Angelegenheit macht.

Da Router die besten Routen im Verhältnis zur Anzahl der bewegten Pakete nur sehr langsam berechnen können, merken sie sich in einer **Routing-Tabelle** die bestmögliche Route zu bestimmten Netzwerken und die dazugehörigen Routing-Metriken.

Routing-Protokolle sorgen für den Austausch von **Routing-Informationen** zwischen den Netzwerken und erlauben es den Routern, ihre Routing-Tabellen dynamisch aufzubauen. Traditionelles IP-Routing bleibt einfach, da **Next-Hop-Routing** benutzt wird. Der Router sendet das Paket an denjenigen Nachbar-Router, von dem er glaubt, dass er am nächsten am Zielnetz liegt. Um den weiteren Weg des Pakets braucht sich der Router nicht zu kümmern. Selbst wenn er falsch lag und das Paket nicht an den "optimalen" Nachbarn gesendet hat, kommt das Paket trotzdem früher oder später am Ziel an.

Obwohl dynamisches Routing sehr komplex werden kann, macht es das Internet sehr flexibel, und erlaubte das exponentielle Wachstum des Internets seit der Einführung von IP im Jahre 1983. Wenn Teile der Backbones ausfallen (so geschehen z.B. im Sommer 2002, als der Carrier KPNQwest sein europaweites Glasfasernetz wegen Insolvenz abschalten musste), können innerhalb von Sekunden Alternativrouten propagiert werden und die betroffenen Netzteile weiträumig umgangen werden.

Dem Ausfall des Standardgateways, das ist meist der erste Router vom Sender aus gesehen, wirkt dynamisches Routing jedoch nicht entgegen. Hierfür wurden HSRP und VRRP entwi-

ckelt, da ein Host im Normalfall keine Alternative zum Standardgateways hat ist dies der wichtigste Router der Route.

Routing-Algorithmen benutzen zwei grundlegende Verfahrensweisen:

- *Teile der Welt mit, wer deine Nachbarn sind*: Link-State-Routing-Protokolle wie z.B. OSPF
- *Teile deinen Nachbarn mit, wie die Welt aussieht*: Distanzvektor-Protokolle wie z.B. das Routing Information Protocol (RIP).

Eine **Routing-Metrik** ist ein Wert, mit dessen Hilfe ein Routing-Algorithmus feststellen kann, ob eine Route im Vergleich zu einer anderen besser ist. Metriken können Informationen wie z.B. Bandbreite, Verzögerung, Hop Count, Pfadkosten, Last, MTU, Verlässlichkeit und Kommunikationskosten berücksichtigen. In der Routing-Tabelle werden nur die bestmöglichen Routen gehalten, während Link-State- oder topologische Datenbanken alle anderen Informationen beinhalten.

Abhängig davon, ob der Router Teil eines autonomen Systems ist oder gar dessen Grenze bildet, verwendet er Routing-Protokolle aus verschiedenen Klassen:

- **Ad hoc Routing-Protokolle** werden in Netzwerken mit wenig oder keiner Infrastruktur verwendet.
- **Interior Gateway Protocols** (**IGP**s) tauschen Routing-Informationen in einem einzelnen autonomen System aus. Häufig verwendet werden:
- IGRP/EIGRP (Interior Gateway Routing Protocol/ Enhanced IGRP)
- OSPF (Open Shortest Path First)
- RIP (Routing Information Protocol)
- IS-IS (Intermediate System to Intermediate System)
- **Exterior Gateway Protocols** (**EGP**s) regeln das Routing zwischen verschiedenen autonomen Systemen. Dazu gehören:
- EGP (mit dem alten Exterior Gateway Protocol wurden früher die Internet-Backbones verbunden. Es ist inzwischen veraltet)
- BGP (Border Gateway Protocol: seit 2002 in der Version BGP4)

Allgemeiner bezeichnet **Routing** auch Navigieren in einem Netz von miteinander verbundenen Knoten. Der kürzeste Weg kann zum Beispiel mit dem Algorithmus von Dijkstra gefunden werden.

PING

ping (*Packet InterNet Groper*) ist ein Computerprogramm um zu überprüfen, ob ein bestimmter Host in einem TCP/IP-Netzwerk erreichbar ist.

Dazu sendet es ein ICMP Echo-Request-Paket an die Zieladresse. Der Empfänger muss laut Protokollspezifikation eine Antwort zurücksenden: ICMP *Echo-Reply*. Ist der Zielrechner

nicht erreichbar, antworten Router: *Network unreachable* (Netzwerk nicht erreichbar) oder *Host unreachable* (Gegenstelle nicht erreichbar).

Einige Parameter sind bei Ping einstellbar. Zum Beispiel bestimmt die Wiederholrate, wie häufig ein Paket gesendet wird. Die Paketgröße bestimmt die Größe des ICMP-Echo-Request-Pakets.

Beispiel

```
# ping de.wikipedia.org

PING de.wikipedia.org (130.94.122.197): 56 data bytes
64 bytes from 130.94.122.197: icmp_seq=0 ttl=239 time=222.1 ms
64 bytes from 130.94.122.197: icmp_seq=1 ttl=239 time=222.5 ms
64 bytes from 130.94.122.197: icmp_seq=2 ttl=239 time=222.4 ms
64 bytes from 130.94.122.197: icmp_seq=3 ttl=239 time=223.1 ms
64 bytes from 130.94.122.197: icmp_seq=4 ttl=239 time=223.7 ms

-- de.wikipedia.org ping statistics --
5 packets transmitted, 5 packets received, 0% packet loss
round-trip min/avg/max = 222.1/222.7/223.7 ms
```

Es werden Datenpakete an den Rechner **de.wikipedia.org** gesandt. Vom Programm wird die Zeit gemessen, bis die Antwort des Hosts eintrifft. Die Zeitangabe sagt aus, wie lange es dauert, bis ein Datenpaket zum Host und wieder zurück braucht. Man kann daran grob erkennen, ob die Gegenstelle funktioniert und mit welcher Verzögerung bei einer Verbindung zu rechnen ist.

Aus einer fehlenden Antwort kann man allerdings nicht eindeutig darauf schließen, dass die Gegenstelle nicht erreichbar ist. Manche Hosts sind nämlich so konfiguriert, dass sie ICMP Pakete ignorieren und verwerfen.

Herkunft des Namens

Der Name ping stammt aus der Kriegstechnik. Seit dem Zweiten Weltkrieg wird Sonar eingesetzt, um U-Boote aufzuspüren. Das dabei ausgestrahlte Schallsignal hört sich in einem U-Boot wie ein hohes Klopfgeräusch an, was lautmalerisch als "Ping" beschrieben wird.

Traceroute

Traceroute ist ein TCP/IP Tool mit dem Informationen darüber ermittelt werden können, welche Computer ein Datenpaket über ein Netzwerk passiert, bis es bei einem bestimmten Host ankommt. Traceroute arbeitet mit dem ICMP-Teil der IP-Spezifikation, über das Rückmeldungen gegeben werden. Dabei werden IP-Pakete mit immer höherem *Time-to-live*

(TTL) Wert gesendet und ICMP Antworten von Rechnern "auf dem Weg" oder dem Ziel-Rechner ausgewertet.

Die Anzeige von traceroute zeigt nicht immer den tatsächlichen Weg, den die Datenpakete nehmen. Es wird beeinflusst von Firewalls, fehlerhaften Implementierungen des IP Stacks, Network Address Translation und anderen Faktoren.

In Windows ist Traceroute als *tracert* implementiert

Beispiel

```
$ traceroute wikipedia.de
traceroute to wikipedia.de (130.94.122.197), 30 hops max, 40 byte packets
 1 fli4l.Netz1 (192.168.0.1) 0.765 ms 0.651 ms 0.497 ms
 2 217.5.98.7 (217.5.98.7) 14.499 ms 14.648 ms 21.394 ms
 3 217.237.152.46 (217.237.152.46) 14.831 ms 13.655 ms 13.403 ms
 4 62.154.14.134 (62.154.14.134) 118.090 ms 119.522 ms 119.665 ms
 5 p16-1-0-3.r20.asbnva01.us.bb.verio.net (129.250.9.141) 117.004 ms 117.370 ms  117.073 ms
 6 p64-0-0-0.r21.asbnva01.us.bb.verio.net (129.250.2.35) 119.105 ms 119.284 ms 119.206 ms
 7 p16-0-1-2.r20.plalca01.us.bb.verio.net (129.250.2.192) 180.035 ms 195.498 ms 178.704 ms
 8 p16-1-0-0.r06.plalca01.us.bb.verio.net (129.250.3.81) 177.280 ms 177.263 ms 176.692 ms
 9 p4-0-3-0.r00.sndgca01.us.bb.verio.net (129.250.3.10) 194.322 ms 193.477 ms 193.743 ms
10 ge-1-1.a03.sndgca01.us.da.verio.net (129.250.27.84) 192.527 ms 193.003 ms 192.464 ms
11 Pliny.wikipedia.org (130.94.122.197) 192.604 ms 193.875 ms 194.254 ms
```

Download

Download (wortwörtlich *hinunter laden*) oder **Herunterladen** ist ein Begriff aus dem IT-Jargon.

Bei einem Download werden Daten von einer Gegenstelle (z. B. Netzrechner, Internet) angefordert und zum Rechner übertragen. Der Download ist somit das Gegenstück zum Upload

Für eine sichere Datenübertragung werden häufig auch von Quelle und Übertragungsergebnis CRC-Prüfsummen gebildet und miteinander verglichen, um eine fehlerfreie Transaktion sicherzustellen.

Neben der Vollständigkeit der Daten ist die Übertragungsgeschwindigkeit ein wichtiges Kriterium. Da technisch bedingt oft die zur Verfügung stehende Übertragungsbandbreite nicht immer voll genutzt werden kann, verwenden bestimmte Download-unterstützende Programme alternative Methoden zur Beschleunigung (z.B. Öffnen mehrerer paralleler virtueller Verbindungen zum Server zur Umgehung der durch den Protokoll-Overhead ausgelösten Latenz).

Spezielle Download-Protokolle umgehen die Nachteile einfacher Client-Server-Verbindungen, indem sich die ladenenden Rechner gegenseitig beim Download unterstützen

und so den Server entlasten. Damit lässt sich der Download häufig erheblich beschleunigen (siehe BitTorrent).

Die aus dem englischen stammende Vokabel wird neben der deutschen Entsprechung "herunterladen" auch vermehrt im Deutschen verwendet (*downloaden, ich loade down, downgeloadeten...*).

UPLOAD

Upload (wortwörtlich *hinauf laden*) oder **Hochladen** ist ein Begriff aus dem IT-Jargon.

Bei einem Upload werden Daten von einem Rechner zur Gegenstelle (z. B. Netzrechner, Mailbox, Internet-Server) übertragen. Der Upload ist somit das Gegenstück zum Download.

Der Unterschied zum Download besteht darin, dass beim Upload der Nutzer seine Daten vom eigenen Computer auf einen fremden Rechner schickt während er beim Download Daten auf den eigenen Rechner holt.

Die Herkunft des Wortes ist etwas unklar. Vielleicht dient hier das Bild eines auf einen Berg oder auf einem Satelliten stehenden, für alle sichtbaren Rechners, zu dem Daten "hoch" geschoben werden als Vorbild.

Ein anderer Ansatz ist jener, dass der eigene Rechner als untergeordnete Einheit im Netzwerk gesehen wird.

Uploads können über viele Netzwerkprotokolle, z. B. HTTP oder FTP erfolgen.

Dienste im Internet

Mailbox

Eine **Mailbox**, englisch **Bulletin Board System** (BBS) genannt (der Name **Mailbox** ist ein Scheinanglizismus), ist ein meist privat betriebenes Rechnersystem, das per DFÜ zur Kommunikation genutzt werden kann. Jeder Benutzer (User) der Mailbox hat ein eigenes Postfach, in dem elektronische Nachrichten für ihn gespeichert und von ihm abgerufen werden können. Zudem gibt es meist öffentliche Bereiche, oft Foren, Bretter oder Echos genannt, in denen die User sich austauschen und diskutieren können. In der Regel bieten Mailboxen darüber hinaus einen Download-Bereich für Dateien an.

Viele Mailboxen sind untereinander vernetzt und tauschen in regelmäßigem Abstand (meist mehrfach am Tag) ihre Daten untereinander aus. Auf diese Weise können User unterschiedlicher Mailboxen schnell und kostengünstig miteinander kommunizieren.

Geschichte

Mit der schnellen Verbreitung der ersten Personalcomputer, vor allem des Apple II, und der ersten brauchbaren Modems, enststanden die privat betriebenen Mailboxen ab ca. 1978/79 vor allem in den Großstädten der USA. Dort waren damals Ortsgespräche kostenlos und kostenlose Ferngrespräche konnten wegen eines kurz zuvor bekanntgewordenen Designfehlers des Telefonnetzes leicht illegal erschlichen werden (z. B. durch Phreaking), was die Verbreitung der Mailboxen v.a. in den Ballungsräumen enorm beschleunigte. In Deutschland, wo die Personalcomputer etwas später aufkamen, 1980 unter Postminister Kurt Gscheidle (SPD) der Zeittakt für Ortsgespräche eingeführt wurde, und das Monopol der Bundespost brauchbare Modems verhinderte, war der Start der Szene viel langsamer und erreichte nie die amerikanischen Ausmaße.

Das weltgrößte private Mailbox-Netz, das FidoNet (kurz Fido) entstand 1984 und verbreitete sich schnell weltweit. Kurze Zeit später entstanden in Deutschland weitere Mailbox-Netze wie das MausNet, Z-Netz, AmNet, die aber keine internationale Verbreitung fanden.

Obwohl die von den verschiedenen Netzen verwendete Software zu einander inkompatibel war, entstanden zwischen diesen Netzen rasch Schnittstellen, so genannte Gateways, mit denen über die Netzgrenzen hinweg Nachrichten verschickt werden konnten.

Mitte der 1990er Jahre erreichte die Zahl der Mailbox-User ihren Höhepunkt. Sie wird auf etwa 1,56 Millionen nur im FidoNet geschätzt. Mit der seitdem zunehmenden Verbreitung des Internet gingen die Benutzerzahlen jedoch stark zurück.

E-Mail

E-Mail (*Electronic Mail*) ist ein Dienst in Computernetzen (vor allem im Internet), der es erlaubt, elektronische Nachrichten zwischen einem Sender und Empfängern auszutauschen (*„E-Post“*). E-Mail ist neben dem WWW der derzeit am häufigsten genutzte Dienst des Internets. Zugleich bezeichnet **E-Mail** die Nachrichten, die man sich in diesem Dienst versenden kann (*„E-Brief“*).

Neben dem Internet ist auch in nahezu allen anderen Computer-Netzen (z. B. ax.25, Novell, Windows, BTX) der E-Mail-Versand nach eigenen Verfahren vorgesehen, seit den 1990er-Jahren aber praktisch bedeutungslos. Lediglich X.400, ein offener, weltweiter Standard, der vor dem Internet definiert wurde, wird heute noch ernsthaft benutzt.

Für *E-Mail* gibt es nur eine Schreibweise, analog zu E-Gitarre, H-Milch oder A-Bombe. *Email*, *EMail* oder *eMail* sind daher falsch.

Format und Aufbau der Adressen

Das Format der Internet-E-Mails wird im RFC 2822 festgelegt und hat folgende Syntax:

benutzer@domain

wobei *benutzer* für eine eindeutige Zeichenkette steht. Diese Zeichenkette darf z. Zt. (*Stand 02.2004*) keine Umlaute und - mit Ausnahme von Punkt, Binde- und Unterstrich - keine Sonderzeichen enthalten. Für *domain* gelten die Syntaxregeln des Domain Name Service.

domain bezeichnet die DNS-Domain, in der sich der *benutzer* befindet. Der Domainname wird dabei durch eine Anfrage an einen Domain Name Server (DNS) aufgelöst und die IP-Adresse des eigentlichen Mailservers durch einen Eintrag "MX" bestimmt.

Benutzer- und Domain werden durch das Zeichen @ (Aussprache engl. "at" oder umgangssprachlich "Klammeraffe", "Affenohr" oder "Affenschwanz") getrennt.

Wichtige Benutzernamen für E-Mailadressen sind die sogenannte *Role-Accounts*. Diese sind z.B.:

- **abuse** für Missbrauchsmeldungen
- **webmaster** um Betreiber einer Webseite zu kontaktieren
- **postmaster** für Probleme betreffend den Mailempfang bzw. -versand
- **hostmaster** bei Zugriffsproblemen oder Probleme bei der Anbindung an das Internet
- **newsmaster** für den Betreuer eines Newsservers

Meist werden Klartext-Nachrichten verschickt. Dabei werden oft Abkürzungen oder Smileys verwendet. Zudem können E-Mails auch mit anderen elektronischen Inhalten wie Bildern, binären Dokumenten oder aufgenommener Sprache als so genannte Dateianhänge verschickt werden.

Der Betreff ist für den Empfänger besonders in Anbetracht der steigenden Anzahl unerwünschter Werbung (Spam) eine wichtige Kurzinformation und sollte daher nie fehlen. Das **Betr.: bzw. Conc:** stellt eine Beziehung des Senders und/ oder des Empfängers zu einer Sache, erforderlichen Aktion oder zu einem Ereignis her.

Verwendete Protokolle

- SMTP ist zum Mailversand und -Transport.
- POP3 ist zum Abruf von Mails von einem Mailserver.
- IMAP dient dazu, auf Mailboxen zuzugreifen, die auf Mailservern liegen.
- UUCP ist ein Protokoll, mit dem E-Mails gesammelt werden und beim nächsten Verbindungsaufbau verschickt werden.
- MIME kodiert E-Mails und legt ihre Struktur fest.

Funktionsweise und Nutzung

E-Mail-Programme

Zur Nutzung von E-Mail braucht man ein E-Mail-Programm, auch *E-Mail-Client* oder *Mail-User-Agent* (*MUA*) genannt. Dieses holt die E-Mails (meist mittels der Protokolle POP3 oder IMAP) vom E-Mail-Server des Providers ab, der sie in einer Mailbox auf dem Server gespeichert hat. Zum Versand einer E-Mail schickt das E-Mail-Programm des Benutzers diese üblicherweise per SMTP (Simple Mail Transfer Protocol) an den E-Mail-Server des Providers (*Smarthost*), der diese dann an den E-Mail-Server des Empfängers weiterschickt. Dort kann sie dann der Empfänger wiederum per POP3 oder IMAP mit seinem E-Mail-Programm abholen. Tauchen beim E-Mail-Versand Probleme auf, z.B. fehlerhafte E-Mail-Adresse oder volle Mailbox des Nutzers so wird die versendete E-Mail "gebounced" (zurückgeschickt), wobei als Absender der so genannte Mailerdaemon auftritt. Ein E-Mail-Programm braucht zur Konfiguration üblicherweise Angaben über das Benutzerkonto (englisch *account*) des Benutzers (Benutzername und Passwort), außerdem die Netzadresse des Servers zum E-Mail-Empfang (z. B. "pop.providername.de") und zum Versenden (z. B. smtp.providername.de).

Webmail

Als alternatives Verfahren zur Verwendung eines E-Mail-Programms hat sich auch die Nutzung von Webmail etabliert. Webmail ermöglicht die Verwaltung von E-Mails mithilfe eines Webbrowsers.

Schwächen

Das herkömmliche E-Mail-System besitzt mehrere Schwächen. Die meisten Nachrichten werden im Klartext verschickt, können also prinzipiell auf jedem Rechner, den die Nachricht auf ihrem Weg vom Absender zum Empfänger passiert, gelesen werden. In der Analogie zur klassischen Post repräsentiert die E-Mail die Postkarte, nicht den durch einen Umschlag vor neugierigen Blicken geschützten Brief. Des Weiteren kann E-Mail i.d.R. mit jedem beliebigen Absender verschickt werden, was oft für unerwünschte Werbung (Spam) benutzt wird. Für beide Probleme (Absenderauthentifizierung und Verschlüsselung) existieren die Verfahren PGP und das freie GnuPG sowie S/MIME, die jedoch noch nicht besonders weit verbreitet sind.

Klammeraffe

Der @ geschriebene Buchstabe, der auch als "**Klammeraffe**" bzw. "Affenschwanz", "Affenohr" oder "Affenschaukel" in der deutschen Umgangssprache bekannt ist, gilt in Deutschland als ein allgemein bekanntes Symbol für das Internet allgemein und E-Mails im speziellen.

Ursprünglich bedeutet "@" im Englischen Geschäftsverkehr dasselbe wie "à" im Deutschen ("4 Zitronen à 20 Cent" <–> "4 lemons @ 20 cents"). Da es im kaufmännischen Bereich in anglophonen Ländern häufig verwendet wurde, gelangte es sehr bald auf Schreibmaschinen und so letztendlich auch auf deutsche Computertastaturen.

Die beiden Schreibweisen "à" und @ nahmen beide ihren Ursprung im Französischen. Im Englischen wird durchgängig die Aussprache *at* (wie in "I'm at home") benutzt und das Zeichen heißt *commercial at*. Das @ wird im Deutschen teilweise auch als "à" oder sogar als "shift-alpha" ausgesprochen.

Als bei der Erfindung der E-Mail im ARPA-Netz (ca. 1971) nach einem noch ungenutzten Zeichen im Schriftsatz amerikanischer Fernschreiber (ASCII) gesucht wurde, das Benutzer- und Rechnername eindeutig trennen sollte, stieß man auf das @ und benutzte es als "at"-Symbol in E-Mail-Adressen. Das Zeichen passte so auch von der Benennung her, bezeichnete der Benutzername doch einen einzelnen Benutzer und der Rechnername ursprünglich meist den Großrechner des Betriebs oder Instituts, wo er arbeitete.

In älteren Programmiersprachen (z. B. BASIC, dBase) wurde @ manchmal auch als Operator für Positionsangaben verwendet, der BASIC-Befehl `PRINT @ 12,10,"HALLO"` stellt zum Beispiel das Wort "HALLO" in der zehnten Zeile und ab der zwölften Spalte des Bildschirms dar.

Die spezielle Assoziation mit dem Internet in Deutschland ergibt sich schlicht daraus, dass das Zeichen hierzulande vor dem Internet-Boom höchstens Programmierern und Heimcomputer-Nutzern bekannt war, da es im deutschen Schriftverkehr traditionell keine Rolle spielte. Im englischen Sprachraum dagegen war das Zeichen bereits vorher lange in Gebrauch, daher existiert eine derartige Assoziation dort nicht.

Seit Anfang 2004 ist der Klammeraffe auch offizieller Bestandteil des Morsecodes.

Mailingliste

Eine **Mailingliste** (ML) bietet einer geschlossenen Gruppe von Menschen Nachrichtenaustausch in Briefform. Dieser Nachrichtenaustausch ist innerhalb der Gruppe öffentlich. Besonders häufig sind Mailinglisten im Internet, wo sie mittels E-Mail realisiert werden.

Mailingliste per E-Mail

Um den anderen Mitgliedern der ML eine Nachricht zukommen zu lassen, muss dafür Sorge getragen werden, dass sie alle Adressaten erreicht. Bei kleinen Gruppen genügt es, für diesen Zweck alle Empfängeradressen durch jeweils ein Komma getrennt in das Empfängerfeld zu schreiben. Bei großen Interessengemeinschaften (wie sie Im Internet üblich sind), schickt man sie an eine zentrale Adresse. Von dort aus wird die Nachricht jedem Mitglied sofort oder einmal täglich als Tageszusammenstellung (Digest) zugesandt.

Mailinglisten sind vergleichbar mit Foren oder dem Usenet. Der Vorteil gegenüber einem Webforum ist, dass die Beiträge offline gelesen und geschrieben werden können. Ein Listenserver (das Programm, das die Nachrichten weiterverteilt) ist viel einfacher einzurichten als eine Newsgroup im Usenet. Außerdem besitzen viel mehr Internetnutzer ein Programm zum Lesen von E-Mails als von Newsgroup-Nachrichten und können es auch bedienen.

Wie auch im Usenet haben sich auf Mailinglisten einige Besonderheiten gegenüber dem normalen E-Mail-Verkehr entwicklet. Ein besonders häufiger Verstoß gegen die Netiquette sind Thread-Hijacking, was bedeutet dass der Benutzer einfach nur zu faul ist für ein neues Thema einen neuen Thread zu starten.

Der Listenserver, der die Emailadressen verwaltet, kann so konfiguriert werden, dass zu einer Mailingliste nur bestimmte Personen zugelassen werden. Auf diese Weise kann man die Kontrolle über die Abonnenten behalten.

Manche Mailinglisten werden von einem Moderator betreut, der eintreffende Mails begutachtet bevor sie an alle Teilnehmer weitergeleitet werden. Der Moderator kann Artikel aussortieren, die rein werbenden Charakter besitzen (Spam), themenfremd sind oder gegen die Netiquette verstoßen.

Häufig genutzte Listensever sind Majordomo, Mailman und ezmlm/idx.

vCard

Eine **vCard** ist eine "**elektronische Visitkarte**" (Dateinamenserweiterung "*.vcf"), die ein Benutzer mit einem Mausklick direkt in das Adressbuch seines E-Mail-Programms übernehmen kann. vCards können in HTML-Seiten eingebettet oder als Dateianhängen zu einer

E-Mail verschickt werden. Mobiltelefone und PDAs verwenden auch vCards, um Kontaktinformationen zu speichern und mit IrDA oder Bluetooth auszutauschen.

Der Inhalt und Aufbau von vCards sind vom Internet Mail Consortium (IMC) standardisiert, und viele E-Mail-Programme können Kontaktdaten als vCard lesen und exportieren. Leider ist die Unterstützung oft unzuverlässig; eine vCard, die mit einem bestimmten Programm erstellt wurde, kann für ein anderes Programm unleserlich sein, Umlaute können verloren gehen, etc. Derzeit bietet das IMC keinen offiziellen Test, mit dem die Qualität einer vCard geprüft werden kann oder eine Liste der Produkte, die vCards (zuverlässig) unterstützen.

GNU Privacy Guard

GnuPG (**Gnu P**rivacy **G**uard, englisch für **GNU-Wächter der Privatsphäre**) ist ein freies Kryptographie-System, das als Ersatz für PGP dienen kann. GnuPG benutzt nur patentfreie Algorithmen und wird unter der GNU General Public License vertrieben. GnuPG läuft unter Linux, Mac OS X und diversen anderen Unix-Varianten sowie unter Microsoft Windows.

Die Entwicklung von GnuPG wird vom Bundesministerium für Wirtschaft und Arbeit (BMWA) und Bundesministerium des Innern (BMI) im Rahmen der Aktion "Sicherheit im Internet" gefördert, um eine frei verfügbare Verschlüsselungssoftware für jedermann zur Verfügung zu stellen. Da der Quellcode jedermann offen steht, hat GnuPG gegenüber dem kommerziellen PGP deutliche Vorteile, z. B. wird sichergestellt, dass man durch GnuPG selbst nicht ausspioniert wird..

GnuPG kann benutzt werden, um Texte (z. B. E-Mails) zu verschlüsseln und digital zu unterschreiben. Dafür werden zwei verschiedene Schlüssel benutzt: Ein privater Schlüssel, auf den nur der Eigentümer Zugriff hat und ein öffentlicher Schlüssel, der auf einem so genannten 'Keyserver' gespeichert werden kann und benötigt wird, um die Korrektheit der digitalen Unterschrift zu verifizieren. Um die Echtheit der öffentlichen Schlüssel sicherzustellen existieren verschiedene Verfahren, bei denen teilweise über mehrstufige Systeme Vertrauen vererbt werden kann. Häufig werden auf Szene-Treffen so genannte 'Keysignpartys' veranstaltet, bei denen öffentliche Schlüssel ausgetauscht werden und die Echtheit durch Vorlage eines Ausweises sichergestellt wird.

Weblinks

- http://www.gnupg.org – GNU Privacy Guard
- (GnuPG) Mini Howto deutsch (http://www.gnupg.org/howtos/de/)
- deutschsprachiges Handbuch (http://www.gnupg.org/gph/de/manual/)

GOPHER

Gopher *(engl. Erdhörnchen)* ist ein Informationsdienst, der über das Internet mit Hilfe eines Gopher-Clients oder einem Webbrowser abgerufen werden kann. Für die Wahl des Namens gibt es 2 Theorien: gopher (Taschenratte), das Maskottchen des Bundesstaates und der Universität von Minnesota oder ,,Go for it`` ab.

GESCHICHTE

Gopher ähnelt dem frühen World Wide Web und wurde 1992 an der Universität von Minnesota entwickelt.

Die Überlegung, die zu Gopher führte, war die umständliche Handhabung von FTP, bei dem man sich einloggen muss und über Konsolenbefehle in Verzeichnisse wechseln musste, um die gewünschte Datei zu finden und herunterladen zu können. Zudem wollte man ein einfach zu administrierendes Informationssystem schaffen, das wenig Ressourcen benötigt.

Mitte der 90er Jahre hatte nahezu jede Organisation, die über einen Internetzugang verfügte, wie zum Beispiel Universitäten, aber auch Regierungen, einen Gopherserver und stellte der Allgemeinheit Informationen aus allen Bereichen zur Verfügung.

Mit dem Aufschwung des WWW jedoch ging die Zeit des so genannten *Gopherspace* zu Ende.

Heute gibt es nur noch sehr wenig Gopherserver und Gopher ist weitgehend unbekannt.

FUNKTIONSWEISE

Gopher baut auf das Gopher-Protokoll auf, das in RFC 1436 definiert ist.

Gopher bietet im Gegensatz zu HTML-Seiten ein automatisch generiertes Menü an, das aus den im aktuellen Verzeichnis befindlichen Dateien generiert wird. Der Gopher-Server erkennt dabei, ob es sich um Verzeichnisse oder Dateien handelt und zeigt dies durch entsprechende Symbole an.

Zusätzlich bieten Gopher-Server auch Konfigurationsdateien an, die es dem Betreiber erlauben, Verweise auf externe Gopherserver zu generieren.

Für den Gopherserver *gopherd*, der zum Beispiel bei der Linux-Distribution Debian mitgeliefert wird, sieht diese Datei im Aufbau folgendermaßen aus:

```
Name=Web Server on Athene
Type=h
Path=GET /
Host=athene.dnsalias.org
Port=80
#
Name=NCT Gopher Server
```

```
Type=1
Port=70
Path=/
Host=gopher.nct.de
```

In dieser Datei ist zum einen ein Verweis auf einen Webserver, aber auch ein Verweis auf einen anderen Gopherserver definiert.

Abgespeichert wird diese Datei in einem Verzeichnis des Gopherservers unter dem Namen . *Links*. Bitte den Punkt vor dem Dateinamen beachten.

Im Gegensatz zu Webseiten sind Gopherseiten reine Textdateien ohne Formatierung, wie zum Beispiel Fettschrift oder eingebettete Grafiken.

Clients

Für Gopher gibt es eigene Clients, die jedoch nicht bei allen Betriebssystem-Distributionen beigelegt werden. Eine Möglichkeit aber, den Gopherspace zu erforschen, bieten Webbrowser, wie zum Beispiel Mozilla. Der Internet Explorer von Microsoft hat diese Fähigkeit vor einiger Zeit aufgrund eines Sicherheitsbugs verloren - man hielt Gopher für nicht wichtig genug, ob das Loch zu schließen. Im WWW findet man außerdem Webseiten, die eine Schnittstelle vom Gopherspace in das WWW bereitstellen.

Um eine Gopherseite mit einem Webbrowser aufzurufen, gibt man die Zeile `gopher://<gopherserver>/` ein. Beispiel: `gopher://gopher.floodgap.com/`

Usenet

Das **Usenet** (*USENET*, urspr. *Unix User Network*) ist ein weltweites elektronisches Netzwerk aus Diskussionsforen, das aus Newsgroups besteht und an dem jeder, der z. B. über das Internet Zugang zu einem Newsserver hat, teilnehmen kann.

Die Funktionsweise des Usenet wird oft mit Schwarzen Brettern, z. B. im Supermarkt, verglichen. Diese Analogie gibt jedoch nur einen Teilaspekt des Usenet wieder, da die Kommunikation über Schwarze Bretter in der Regel nur in eine Richtung (simplex bzw. undirektional) läuft. Ein passenderer Vergleich, und daher rührt auch die Usenet-Sprache, ist das Zeitungswesen: Jemand schreibt einen Artikel (news, article) für die Zeitung (newsgroup). Ein Leser nimmt auf diesen Artikel Bezug und schreibt einen Leserbrief, den er an die Redaktion schickt. Durch die Veröffentlichung wird dieser Leserbrief wieder zu einem Artikel, auf den sich ein weiterer Leser beziehen kann usw...

Das Usenet unterscheidet sich darin, dass es keine Redaktion hat, die eine Vorauswahl der zu veröffentlichenden Artikel oder Leserbriefe trifft. Weitere Vorteile sind die Geschwindigkeit und die hohe Teilnehmerzahl. Innerhalb weniger Stunden können zu kontroversen Themen riesige Diskussionsbäume (Threads) entstehen.

Im Gegensatz zum Chat kann man im Usenet nicht schon nach ein paar Sekunden Antwort erwarten, denn die Nachricht muss zunächst von Server zu Server weitergereicht werden. Außerdem lesen viele Teilnehmer die Beiträge offline, d. h. sie laden sich die neuen Beiträge der von ihnen abonnierten Gruppen ein- oder mehrmals am Tag lokal runter, schreiben Antworten offline und senden diese erst dann gesammelt zurück an den Server. Auch sind die Umgangsformen verglichen zum Chat meist viel 'geregelter' (Netiquette).

Newsgroups

Um das Usenet übersichtlich zu gestalten wird es in einzelne Newsgroups unterteilt. Das sind Gruppen, in denen nur über ein bestimmtes Thema (Topic) diskutiert wird. Zum Beispiel über Festplatten. Oder Kinofilme. Oder elektronische Musik. Im Prinzip gibt es zu jedem Thema eine passende Gruppe. Newsgroups sind hierarchisch gegliedert, also zum Beispiel so:

- de.rec.alpinismus (news:de.rec.alpinismus)
- de.rec.buecher (news:de.rec.buecher)
- de.rec.misc (news:de.rec.misc)
- de.rec.musik.klassik (news:de.rec.musik.klassik)
- de.rec.musik.machen (news:de.rec.musik.machen)
- de.rec.musik.misc (news:de.rec.musik.misc)
- de.sci.chemie (news:de.sci.chemie)
- de.soc.menschenrechte (news:de.soc.menschenrechte)

de steht für den deutschsprachigen Teil des Usenet. rec (von "recreation") steht für Freizeitthemen (im weitesten Sinne), sci (von "science") für die Wissenschaft, soc für Soziales. In den misc-Gruppen (von "miscellaneous") landen die Themen, die in den Untergruppen keinen Platz finden.

Für Anfänger (Newbies) besonders empfehlenswert sind die Gruppen der Unterhierarchie de.newusers

Geschichte

Das Usenet entstand Ende der 70er-Jahre in den USA als Verbindung zweier UNIX-Rechner an der University of North Carolina und der Duke University. Der Datenaustausch erfolgte über herkömmliche Telefonleitungen mit dem UNIX-Protokoll UUCP (UNIX To UNIX Copy).

Schon bald wurden weitere Rechner in das Netz integriert, wegen des verwendeten UUCP-Protokolls war das Netz jedoch auf UNIX-Rechner beschränkt. Dort rührt auch der Name Usenet her, von *UNIX User Network*.

Über UUCP bestand die Möglichkeit, zum einen persönliche Nachrichten auszutauschen (E-Mail), zum anderen in öffentlichen Foren teilzunehmen.

Um einen besseren Überblick über die verfügbaren Newsgroups zu haben, wurden diese hierarchisch nach sieben Hauptthemen unterteilt, (die so genannten *Major Seven* oder *Big Seven*). Diese waren (und sind):

- **comp:** Themen rund um den Computer
- **sci:** Wissenschaft und Technik ("science")
- **soc:** Gesellschaftlichen Themen ("social")
- **talk:** Allgemeine Gespräche über dies und das
- **rec:** Alle Themen rund um Freizeit und Erholung, z. T. auch Kunst und Kultur
- **news:** In dieser Hierarchie ist das Usenet selbst Gesprächsthema
- **misc:** Alles was nicht in einer der oben genannten Newsgroups Thema ist

Aufgrund der technischen Struktur des Usenet blieben dies lange Zeit die einzigen Hierarchien. Das Netz war bis zu dem Zeitpunkt zwar auf einige tausend Rechner angewachsen, der Datenverkehr lief jedoch großteils über wenige zentrale Rechner (Backbones), deren Administratoren ziemlich viel Macht bei der Einrichtung neuer Gruppen hatten.

Dies änderte sich etwa Mitte der 1980er mit Veröffentlichung des Protokolls **NNTP** (Network News Transport Protocoll). NNTP wurde für den Betrieb über TCP/IP-Leitungen entwickelt. Damit konnte der Datenaustausch erfolgreich über das Internet abgewickelt und das Usenet so dezentralisiert werden, denn über das Internet ist prinzipiell jeder Newsserver von jedem Ort aus ansprechbar. Mehr noch: Jeder Administrator kann über seinen eigenen Newsserver eigene Gruppen einrichten und diese anderen Servern zur Verfügung stellen. So entstanden weitere Hierarchien.

Mit der zunehmenden Verbreitung des Usenet außerhalb der USA entstand auch der Bedarf an Newsgroups in anderen Sprachen. So entstand Ende der 1980er / Anfang der 1990er die deutschsprachige Usenet-Hierarchie "de.*" aus der Verschmelzung der deutschsprachigen Hierachien "dnet.*" und "sub.*". Andere Regionen richteten ebenfalls eigene Hierarchien ein. Aber auch Computerfirmen hatten längst die Möglichkeiten des Usenet als Support- und Informationsmedium entdeckt und bauten eigene Newsserver mit eigenen Hierarchien auf, die zum Teil von anderen Servern geführt werden.

Nennenswerte andere Hierarchien:

- **alt:** Die alt.* Hierarchie ist der etwas anarchistische Teil des Usenet. Die Einrichtung neuer Gruppen kann hier relativ formlos erfolgen, dementsprechend viele (aber qualitativ sehr durchsetzte) Newsgroups gibt es hier.
- **alt.binaries:** Dieser Unterhierarchie gebührt nochmals gesonderte Beachtung, da in hier angesiedelten Gruppen auch Postings mit Dateianhängen (Binaries) erlaubt sind. Leider werden diese Gruppen aufgrund des hohen Traffics und teilweise illegaler oder pornographischer Inhalte fast ausnahmslos nur von kommerziellen Newsservern geführt.
- **de:** Der deutsprachige Zweig des Usenet

- **de.answers:** Hier werden regelmäßig FAQs verschiedener Newsgroups gepostet.
- **de.comp:** Computerbezogene Themen
- **de.sci:** Wissenschaftliche und technische Gruppen

Eine interessante Einrichtung war das unabhängige *DejaNews*-Archiv (Deja.com), das es sich zum Ziel gesetzt hatte, die News-Beiträge dauerhaft zu archivieren. Das Archiv von Deja.com reichte bis etwa 1995 zurück. Nach dem Konkurs von Deja.com wurde dessen Datenbestand Anfang 2001 von Google aufgekauft und in das eigene *GoogleGroups* integriert. Ende 2001 wurden von Google ca. 700 Millionen weitere Artikel in das Archiv integriert, die bis in die Anfangszeit des Usenet zurück reichen. *GoogleGroups* ist unter http://groups.google.com zu finden, es bietet allerdings keinen Zugriff auf binary-Gruppen.

Heutzutage kann niemand sagen, wieviele Newsserver und Newsgroups es weltweit gibt. Schätzungen gehen von Zahlen zwischen 50.000 und 100.000 aus. Allein der Newsserver der Freien Universität Berlin führt ca. 20.000 Newsgroups, allerdings fehlt hier unter anderem die sehr umfangreiche alt.binaries Hierarchie komplett.

Technik

Streng genommen ist das Usenet kein Bestandteil des Internet, auch wenn Verbreitung und Zugriff heute weitgehend darüber erfolgen.

Newsserver transportieren die Nachrichten. Verwendete Protokolle: NNTP, UUCP

Teilnehmer des Usenets lesen und schreiben die Nachrichten in den Newsgroups mit einem Newsreader. Zur Kodierung der Nachrichten wird MIME verwendet.

Weblinks

- www.usenet-abc.de (http://www.usenet-abc.de/) - weitere Informationen zur Teilnahme am Usenet
- Einsteiger können über die WWW-Schnittstelle Arcor WebNews (http://www.arcor.de/webnews/) erste Erfahrungen sammeln, oder sich gleich einen Newsreader (http://www.thomas-huehn.de/usenet/newsreaderFAQ.txt) und einen Zugang zum Newsserver der FU Berlin news.individual.de (http://news.individual.de/) besorgen. Dieser ist sehr gut gepflegt und darf zudem nach Anmeldung kostenlos genutzt werden.
- In alten Usenet-Beiträgen kann man bei Google (http://groups.google.com) suchen. Das Archiv umfasst mehr als 20 Jahre der Usenet-Geschichte

TOFU

TOFU ist das Akronym für "Text oben, Fullquote unten" und wird auch als "Top-posting" bezeichnet.

Als **TOFU** werden Beiträge in Newsgroups, Internetforen oder auch ganz allgemein E-Mails bezeichnet, in denen sich der Antworttext oberhalb der komplett zitierten ursprünglichen Nachricht (der oft auch mit dem englischen Begriff *Quoting* umschrieben wird) befindet, auf die sich die Antwort bezieht.

Beispiel

TOFU	in Zitat-Form	ganz ohne TOFU oder Zitat
Das stimmt doch gar nicht. Richtig ist 4. Gruß Max Oskar schrieb: > 2 + 2 = 5 > Wie jedermann weiß > Beste Grüße > Oskar	Oskar schrieb: > 2 + 2 = 5 Das stimmt doch gar nicht. Richtig ist 4. Gruß Max	Lieber Oskar, 2 + 2 = 4 und nicht wie von dir behauptet 5. Gruß Max

Kritik an TOFU

Durch TOFU wächst die Länge einer E-Mail im Verlauf einer Diskussion schnell an und auch die Länge der einzelnen Zeilen überschreitet schnell das vom papierenen Briefverkehr übernommene Limit von 70-80 Zeichen pro Zeile. Dadurch entstehen schnell auch Kammquotings die nur schwer lesbar sind.

Bei längeren E-Mails und Diskussionen hindert TOFU die Leser daran, den Überblick zu bewahren. Der Empfänger einer TOFU-Mail kann nur durch intensive Beschäftigung mit dem kompletten Kommentarbaum herausfinden, auf welche bisherigen Absätze der Schreiber antwortet.

Da die TOFU-Methode im herkömmlichen papiernen Briefverkehr nicht vorkommt wird desweiteren kritisiert, dass TOFU ein sichtbarer Ausdruck der Faulheit des Schreibers sei.

Weite Verbreitung von TOFU im E-Mail-Verkehr

Die starke Präferenz für das Platzieren der Antwort unterhalb des Zitats bzw. die heftige Kritik an TOFU ist in hohem Maße subkulturspezifisch für das Usenet. So ist z. B. im geschäftlichen E-Mail-Verkehr TOFU die übliche Form der Antwort. Da der Usenet-Stil in

vielen Unternehmen als eher ungewöhnlich gilt, werden gelegentliche "Kommentare im Zitat" in einem Einleitungssatz oberhalb des Fullquotes angekündigt, um die TOFU-gewohnte Leserschaft nicht zu verwirren.

Im Gegensatz zum Usenet kann im E-Mail-Verkehr das isolierte Zitieren und Kommentieren von einzelnen Sätzen als unhöflich empfunden werden, ähnlich wie dies in einem papiernen Geschäftsbrief der Fall wäre. Die Vergrößerung von Mails durch mehrere zitierte Vorgängerschreiben spielt im Bewusstsein von Usenet-Nichtnutzern fast keine Rolle, da die anhängenden Zitate fast niemals so umfangreich sind, dass dadurch der E-Mail-Versand über das typische Computer-Netzwerk einer Firma behindert würde.

Im E-Mail-Verkehr erlangte TOFU vor allem durch die Programme von Microsoft unter denjenigen, die keinen Kontakt mit dem Usenet haben, Standard-Status. Zum Teil ist es bei den Microsoft-Programmen fast unmöglich, nicht TOFU zu verwenden.

Alternativen zu TOFU

Vor allem im Usenet und in technischen Mailinglisten wird nur das zitiert, worauf Bezug genommen wird (*full* quote also vermieden), und zweitens die Antwort unterhalb des Zitates platziert.

Neben diesem Zitierstil ist auch die im papiernen Briefverkehr übliche Methode eine Alternative. Zitate aus der vorherigen Nachricht werden dabei nur selten benutzt und stattdessem mit der bewussten Verwendung einiger Schlüsselwörter aus der letzten Nachricht dem Empfänger der Zusammenhang zwischen Antwort und Ursprungsnachricht besser verdeutlicht. Auch das heute schon fast vergessene lateinische *ad* (zum Thema ...) lässt sich verwenden.

Weblinks

- http://www.afaik.de/usenet/faq/zitieren/ - Wie zitiere ich im Usenet?
- textkritik.de - über Verwahrlosung von E-Mails

Webfoum

Ein **Webforum** ist ein Diskussionsforum auf einer Website. Es ist eine Alternative zu den bereits älteren Medien wie Usenet oder früher gebräuchlicher Bulletin Board Systems sowie Mailinglisten.

Üblicherweise besitzt ein Webforum ein bestimmtes Thema und ist nochmals in Unterforen bzw. Unterthemen unterteilt. Im Gegensatz zum Chat erfolgt die Kommunikation nicht in Echtzeit, sondern asynchron. Es können Postings (Diskussionsbeiträge) hinterlassen werden, die dann später von Interessierten gelesen und beantwortet werden können. Mehrere Pos-

tings zum selben Thema werden wie im Usenet zusammenfassend als *Thread* (Faden) oder *Topic* (Thema) bezeichnet.

Man kann Forensysteme nach ihrer Strukturierung der Beiträge kategorisieren in *"threaded-view"*-Foren und *"flat-style"* (bzw. *"linear-style"*) -Foren. Bei Threaded-View-Foren werden die Beziehungen zwischen den Beiträge innerhalb eines Themas in Form eines Baumes dargestellt. Es entsteht eine hierarchische Struktur, durch die man erkennt, welcher Beitrag als Antwort auf welchen anderen Beitrag erstellt wurde. Flat-Style-Foren zeigen alle Beiträge innerhalb eines Themas nach strikter, chronologischer Reihenfolge ihres Erstellungsdatums an.

Es gibt im Gegensatz zum offenen Usenet weitaus weniger offene Foren, die meisten Webforen setzen eine Registrierung voraus. Viele Foren bieten registrierten Benutzern die Möglichkeit, Threads zu "abonnieren", das heißt sich per E-Mail benachrichtigen zu lassen, wenn einem Thread ein neues Posting hinzugefügt wird.

Viele Foren unterscheiden zwischen verschiedenen Benutzerrollen, zum Beispiel gewöhnlichen Forennutzern und Administratoren. Administratoren haben das Recht, Beiträge anderer zu löschen oder zu editieren. Sie können außerdem oft Threads schließen oder sperren, d. h. das Hinzufügen weiterer Diskussionsbeiträge zu einem Thema verhindern.

Internet Relay Chat

Der **IRC** (**Internet Relay Chat**) bietet weltweit die Möglichkeiten zur Echtzeitkommunikation über das Internet. Es ist ein verteiltes Netzwerk, das seine Daten über dezentral verteilte Server, die untereinander verknüpft sind, austauscht. Es gibt mehrere voneinander unabhängige Netze mit Namen wie EFnet, IRCNet, Undernet, DALnet, Freenode oder NewNet. Das Chatsystem ist textbasiert, erlaubt jedoch über Spezialkommandos auch den Austausch von Dateien und sonstigen Informationen. Um sich mit dem IRC zu verbinden benötigt der Benutzer ein Client-Programm.

Gängige IRC-Clients sind z. B.:

- Unix/Linux: ircII, irssi, BitchX, Kopete, Ksirc
- Windows: mIRC, PIRCH
- Macintosh: Ircle, colloquy, aqua
- auf allen drei Systemen: ChatZilla, XChat

Diese Programme bieten meistens bereits eine Auswahl an Servern an, mit denen man sich verbinden lassen kann. Nachdem man mit einem Server verbunden ist, kann man einen oder mehrere Räume (Channels) auswählen, in denen man chatten möchte. Der Name der Channels deutet meistens bereits auf die Vorlieben der sich darin aufhaltenden Personen hin. Ein Channel mit dem Namen #Berlin wird z. B. größtenteils von Berlinern besiedelt sein. Wenn ein Name für einen noch nicht vorhandenen Channel angegeben wird, so wird dadurch ein neuer Channel geschaffen, der solange besteht, bis sich keine Benutzer mehr darin

befinden. (Je nach verwendetem Netz kann ein Channel ggf. auch registriert werden, um Missbrauch des Namens und/oder "Machtübernahmen" zu verhindern.)

Weblinks

- offizielle Webseiten des deutschen IRC-Net (http://irc.fu-berlin.de/) mit kurzer Einführung und weiteren Links
- de.comm.chatsystems FAQ (http://irc.fu-berlin.de/de-comm-chatsystems-faq.html) - Fragen und Antworten zu IRC
- www.irchelp.org (http://www.irchelp.org/) Umfangreiches Informationsarchiv rund um IRC. Auch die spezifischen RFC Dokumente sind hier abrufbar. (Englisch)

Chat

Chat (von engl. to chat plaudern) ist die Bezeichnung für die innerhalb des Internet weit verbreitete Art der direkten Unterhaltung zwischen zwei oder mehreren Personen in Echtzeit. Es ist eine Art Computerkonferenz, die meist allerdings ohne Bilder auskommen muss. Stattdessen gebrauchen Teilnehmer, die zusammen chatten, daher neben dem geschriebenen Wort auch Ersatzbilder (Avatare, Emoticons).

Oft wird in themenbezogenen Chaträumen gechattet. Fast immer besteht auch die Möglichkeit, zu zweit in einen privaten Chatraum zu wechseln.

Es gibt jedoch auch Software-Hersteller die Chat-Programme anbieten welche den Austausch von Audio- und Video-Nachrichten ermöglichen. Dies gilt beispielsweise für die Software iChat des Herstellers Apple Computer.

Um die Kommunikation in einem Chat zu vereinfachen und zu beschleunigen, hat sich ein Chat-Slang etabliert, der vor allem aus Abkürzungen und **Emoticons**, auch als Smilies bekannt, besteht.

Zu beachten ist auch die Chatiquette. Das sind Regeln für die Umgangsformen in einem Chat. Da das Internet zwar **anonym** ist, sollte man doch gewisse **Umgangsregeln** einhalten - genauso wie im realen Leben. Allgemeine Regeln für die Umgangsformen im Internet nennt man **Netiquette**.

Instant Messaging

Instant Messaging (*IM*) ist ein Dienst, der es erlaubt, in Echtzeit zu chatten oder kurze Nachrichten (im *push*-Verfahren) an andere Teilnehmer zu schicken oder diesen kleinere Dateien zukommen zu lassen.

Die meisten Client-Programme (*Instant Messengers*) ermöglichen es, so genannte Buddy-Listen zu erstellen. Dabei werden die Adressen von anderen Teilnehmern gespeichert und es wird gemeldet, sobald diese im Internet sind und ebenfalls IM nutzen.

Die meisten IM-Dienste sind aufgrund verschiedener Protokolle untereinander inkompatibel.

Dem Betreiber des AOL Instant Messengers (AIM) und seit 1998 auch von ICQ, AOL, wurde im September 2002 ein US-Software-Patent auf Instant Messaging zugesprochen.

Mittlerweile gibt es auch universelle IM-Software, die mehrere Protokolle beherrscht, so z. B. Trillian oder Miranda. Jabber unterstützt die Verbindung zu anderen Protokollen serverseitig, so dass am Client nur Jabber benötigt wird.

Weblinks

Bekannnte IM-Dienste:

- AIM: http://www.aol.de/aim
- ICQ: http://www.icq.com
- Jabber: http://www.jabber.org
- Microsoft MSN-Messenger: http://messenger.msn.com
- Yahoo!-Messenger: http://de.messenger.yahoo.com

Andere IM-Software (Multiprotokollclients):

- Bitlbee für Linux: http://www.bitlbee.org (kann ICQ, AIM, YAHOO MSN, Jabber, und wird per IRC-Client angesprochen)
- Kopete vom KDE-Projekt für Linux: http://kopete.kde.org (kann ICQ, AIM, YAHOO MSN, Jabber, IRC, Windows LANs, GaduGadu und SMS)
- Miranda für Windows: http://www.miranda-im.org/ (kann ICQ, AIM, YIM, MSN, Jabber, IRC, QQ, etc.)

IP-Telefonie

IP-Telefonie, auch als **Voice over IP** (kurz **VoIP**) bekannt, ist das Telefonieren über ein Netzwerk auf der Grundlage des Internet Protocol.

Um die IP-Telefonie zu nutzen, gibt es zwei Varianten:

1. Die Verwendung eines Headsets oder eines speziellen Telefonhörers, die an einen Computer angeschlossen werden. Außerdem wird ein Programm benötigt, das die Anwahl des gewünschten Kommunikationspartners übernimmt.

2. Die Verwendung eines speziellen IP-Telefons oder Adapters, das sich als unabhängiger Client in ein Netzwerk einfügt. Hierzu können kabelgebundene oder kabellose (WLAN) Verbindungen genutzt werden.

Die IP-Telefonie kann folgende Wege einschlagen:

- vom Internet zum Festnetz: dazu wird ein Vermittlungsrechner von Internetanbietern benutzt; meist kostenpflichtig.
- vom Festnetz zum Internet: man ruft einen Vermittlungsrechner an, der den Anruf weiterleitet, sofern der empfangende PC empfangsbereit ist.
- vom Internet zum Internet: derzeit über viele proprietäre Protokolle oder Standards wie H.323 realisiert.

Viele Instant Messenger wie zum Beispiel iChat verschmelzen immer mehr zu IP-Telefonie- und Bildtelefonie-Angeboten (dazu benötigt zusätzlich eine Webcam). Leider benutzen die Anbieter dazu noch oft eigene Formate. Auch einige E-Mail-Anbieter ermöglichen IP-Telefonie.

Funktionsprinzip

- Die Sprache wird durch ein Mikrofon in elektrische Signale umgewandelt. Diese werden mit einem Analog-Digital-Wandler und einem Codec in einen digitalen Datenstrom verwandelt.
- Der digitale Datenstrom wird aufgeteilt und in IP-Pakete verpackt.
- Die IP-Pakete werden über ein Netzwerk übertragen.
- Die ankommenden Pakete werden wieder zu einem digitalen Datenstrom zusammengefasst.
- Ein Codec und ein Digital-Analog-Wandler wandeln den Datenstrom wieder in ein analoges Signal um. Ein Lautsprecher wandelt dieses Signal in Sprache zurück.

Um Telefonieren zu können, muss den Teilnehmern das gleichzeitige Sprechen und Hören ermöglicht werden. Der wesentliche Unterschied zu herkömmlicher digitaler Telefonie (ISDN) besteht darin, das die Daten über ein Netzwerk und nicht eine geschaltete Verbindung in einem Telefonnetz übertragen werden. Bei dem Netzwerk kann es sich um ein LAN, WAN oder das Internet handeln. Die IP-Pakete enthalten zwar Audiodaten, unterliegen aber den üblichen Regeln in Netzwerk: Sie können über verschiedene Wege laufen und werden von Routern zum Ziel geleitet.

Technik

Für eine Verbindung zwischen den Teilnehmern müssen die IP-Adressen bekannt sein. Da die meisten Rechner keine statischen IP-Adressen verwenden, werden die aktuellen IP-Adressen über einen Server ausgetauscht, an dem sich die Teilnehmer anmelden. Ist das nicht möglich, müssen die IP-Adressen über eine bekannte Verbindung (Email, Telefon ...) ausgetauscht und manuell eingetragen werden.

Die Sprachinformationen werden von den Endgeräten über Codecs in IP-Pakete umgewandelt. Dabei werden die Daten unterschiedlich stark komprimiert, was zu unterschiedlichen Bandbreiten bei der Übertragung führt. Je nach Codec variiert auch die Sprachqualität. Auch durch die Verzögerungen, die beim Transfer von IP-Paketen bei entsprechendem Verkehr im Netz auftreten, wird die Sprachqualität beeinträchtigt, was im Extremfall zur völligen Unverständlichkeit führt.

Eine Weiterentwicklung des Internet soll die notwendige Übertragungsqualität zur Verfügung stellen, damit die IP-Telefonie in dieser Hinsicht konkurrenzfähig wird zur herkömmlichen Telefonie. Dazu muß aber die Mehrzahl der Router, Server, Multiplexer, etc. die heute das Internet bilden, gegen neue Geräte ausgetauscht werden, was nicht unbeträchtliche Kosten verursachen wird. Die einschlägigen Standardisierungsgremien (IETF, ITU, ETSI,...) arbeiten an Festlegungen, die von allen Herstellern eingehalten werden sollen, damit die Geräte auch kompatibel werden. Bis dahin sind nur Insellösungen möglich, wobei Inseln unterschiedlicher Technik über Gateways miteinander verbunden werden können.

Populäre Codecs

(benötigte netto-Bandbreite in Klammern)

- G.711a bzw. G.711u - ITU-T-Standard (64kbit/s)
- GSM - (13,2kbit/s)
- G.723 - (5,4kbit/s oder 6,3kbit/s)
- G.729 - (8kbit/s)
- iLBC - Internet Low Bandwith Codec - IETF draft (13,9kbit/s)
- SpeeX - Teil des Xiph.org-Projekts (variable Bitraten)

Populäre Signalisierungsprotokolle

Der Rufauf- und abbau erfolgt über ein von der Sprachkommunikation getrenntes Protokoll. Auch die Aushandlung der Parameter für die Sprachübertragung erfolgt über diese Protokolle. Populäre Signalisierungsprotokolle sind:

- SIP - Session Initiation Protocol, IETF RFC 3261
- H.323 - eine ITU-T Empfehlung

Vorteile der IP-Telefonie

Niedrigere Verbindungpreise möglich

Da die Benutzung von IP-Netzen preiswerter ist als die Verwendung von Telefonnetzen, kann die Verwendung der IP-Telefonie Kosten sparen. Dabei muss ein möglichst großer Teil der Strecke über eine IP-Verbindung abgewickelt werden. Enorme Kostenvorteile entstehen beispielsweise bei Auslandsgesprächen oder wenn beide Teilnehmer IP-Netze nutzen. Da aber oft ein Übergang ins herkömmliche Telefonnetz nötig ist, ergeben sich nicht

immer Kostenvorteile. Die Benutzung des öffentlichen Telefonnetzes muß beim Betreiber des Gateways bezahlt werden. Die Gesamtkosten berechnen sich dann aus der Summe der Entgelte, die für die Benutzung der IP-Verbindung und der Telefonverbindung zu zahlen sind.

Einfachere Verkabelung in Firmen

Viele Firmen haben bisher zwei verschiedene Kommunikationseinrichtungen: das Telefonnetz zur Sprachübertragung und das Computernetzwerk zur Datenübertragung. Dadurch ist eine doppelte Verkabelung notwendig und beide Systeme müssen separat gewartet werden. Bei einer Lösung mit IP-Telefonie kann die Telefonverkabelung entfallen. Dafür muss ein Übergang vom Netzwerk ins öffentliche Telefonnetz geschaffen werden. Auch ein schrittweiser Übergang ist möglich, indem nur neue Arbeitsplätze mit IP-Telefonen ausstattet werden.

Zusätzliche Dienste im Telefon

Beim Einsatz von Telefonen mit LC-Display können diese weitere Informationen auf Abruf anzeigen, z. B. kann der aktuelle Kantinenplan online zur Verfügung gestellt werden oder ein konzernweites Telefonbuch abgefragt werden.

Integration von Computer und Telefon

Bei der Integration von IP-Telefonie und Computer werden die Bedienelemente eines Telefons durch ein Programm nachgebildet. Die Benutzeroberfläche kann oft den individuellen Bedürfnissen angepaßt werden. Telefonnummern können direkt aus den Anwendungen (E-Mail-Programm, Adressbuch ...) heraus gewählt werden. Das E-Mail-Programm kann durch Verwendung von VoiceMail als Anrufbeantworter genutzt werden.

Nachteile der IP-Telefonie

Keine gesicherte Übertragungsqualität

Da das Internet paketorientiert ist und Daten auf verschiedensten Wegen überträgt, aber nicht immer schnell genug ist, kann es zu Übertragungsverlusten oder Verzerrungen kommen. Um den Übertraungsverlusten oder Verzerrungen entgegenzuwirken, gibt es Router mit VoIP-Priorisierung und auch entsprechende Software.

Kein Rufnummernplan

Es gibt keinen internationalen oder auch nur nationalen Rufnummernplan, ebensowenig übergreifende Auskunftsdienste.

Entwicklung

Nach mehreren Anläufen Ende der 90er Jahre, in denen sich die IP-Telefonie nicht durchsetzten konnte, scheinen nunmehr mit der Verbreitung von Internet-Verbindungen mit hoher

Bandbreite, z. B. DSL, und dem Vorliegen internationaler Standards die Voraussetzungen für einen erfolgreichen Einsatz gegeben zu sein. Einige Anbieter stellen hierfür kostenlose Software zur Verfügung, Netzbetreiber bieten Services ohne oder mit niedriger Grundgebühr an.

In jüngster Zeit (2003/4) sind Programme auf den Markt gekommen, die durch die Nutzung des SIP-Standards (Session Initiation Protocol) in Verbindung mit Dienstleistungsunternehmen die Brücke zwischen IP-Telefonie und Festnetz schlagen können. Gratis angeboten wird zur Zeit das Programm x-lite.

Schon seit einigen Monaten ist die Gratissoftware Skype erhältlich, mit welcher man mit erstaunlicher Qualität - laut den Entwicklern besser als Festnetzqualität - über das Internet telefonieren kann. Die Software wurde von den gleichen Leuten entwickelt, welche auch schon KaZaA programmiert haben und basiert ebenfalls auf der P2P Technologie. Im Moment sind leider nur Anrufe von PC zu PC möglich, doch noch in diesem Jahr soll es möglich werden über Skype auch Festnetztelefone zu erreichen.

Zukunft

T-Com, die Festnetzsparte der Deutschen Telekom, plant, ihr Telefonnetz bereits *"deutlich vor dem Jahr 2020"* aus Kostengründen vollständig auf Internet-Übertragungstechnik umstellen.

Peer-to-Peer

Peer-to-Peer (engl. *peer* "Gleichgestellter", "Ebenbürtiger" oder "Altersgenosse/in") ist eine verbreitete Lesart für **P2P** und bezeichnet Kommunikation unter Gleichen. Andere Interpretationen von P2P lauten Person-to-Person (Betonung der computergestützten zwischenmenschlichen Kommunikation) und Program-to-Program (Kommunikation zwischen "intelligenten" Agenten).

In der Informationstechnik ist der Gegensatz zum Peer-to-Peer-Prinzip das Client-Server-Prinzip. Hier gibt es den Server, der einen Dienst anbietet, und den Client, der diesen Dienst nutzt. Der Kern des Begriffes liegt in der Rollenverteilung: Wenn man im Internet surft, ist der Browser kein Webserver und mit dem Webserver kann man nicht browsen.

In Peer-to-Peer-Systemen ist diese Rollenverteilung aufgehoben. Jeder Host in einem Computernetzwerk ist ein *peer*, denn er kann Client und Server gleichzeitig sein.

Verwendung

Peer-to-Peer-Strukturen eignen sich immer dann, wenn die Beziehung der Teilnehmer untereinander einem Netz ähnlicher ist als einer Stern-Topologie. Die E-Mail im Internet hat eine Peer-to-Peer-Struktur. Die Mailserver tauschen direkt untereinander die E-Mail aus, man kann das auch eine n:m-Beziehnung nennen.

Ein Webserver stellt dagegen eine Client-Server-Struktur dar: Ein Server hat n Clients, eine 1:n-Beziehung.

Sehr ins Gerede gekommen ist P2P durch Tauschbörsen, wo *jeder mit jedem* Dateien/Daten austauschen kann. Die Kommunikation kann hier entweder über einen Server erfolgen oder völlig dezentral sein. Der Vorteil von letzterem ist, dass es keine einzelne Einrichtung gibt auf die das gesamte Netzwerk angewiesen ist und es deshalb auch nahezu unmöglich ist das Netzwerk in seiner Funktion zu unterbrechen. Das Netzwerk ist auf diese Weise ausfallsicherer und wesentlich schwerer zu überwachen.

Seit kurzem wird die P2P Technologie auch für die Internettelefonie (VoIP) genutzt. Die Entwickler von KaZaA haben eine Gratissoftware entwickelt, welche Anrufe von PC zu PC in erstaunlicher Qualität ermöglicht.

Manche (moderne) Groupware-Syteme bauen ebenfalls auf dem P2P-Konzept auf. Eines der bekanntesten ist Groove Workspace von Groove Networks.

STANDARDISIERUNG

Die Zukunft der P2P-Technologie wird vor allem davon abhängen, ob es gelingt, einen Standard zu definieren - eine Art Plattform-Technologie, die es ermöglicht, weitere Anwendungen aufzusetzen. JXTA ist ein solcher Standard, der stark von SUN unterstützt wird und Open Source ist. SUN stellte die zurzeit umfangreichste und stabilste Referenzimplementierung her.

Im Zusammenhang mit dem Internet sind **Tauschbörsen** Computernetzwerke, in denen es möglich ist, Dateien über das Internet anzubieten oder herunterzuladen. Die meisten Tauschbörsen besitzen keinerlei Inhalts- bzw Copyright-Kontrollen, sodass Inhalte (meist aktuelle Musik im MP3-Format oder Kinofilme) auch illegal getauscht werden können.

Der Begriff wurde sehr stark durch die Medien geprägt, richtiger wär die Bezeichnung *Kopierbörse*, weil die Daten über Netze weiterkopiert werden, ohne dass das Original selbst den Besitzer wechselt. Aus diesem Grund sind auch Verwertungsgesellschaften für Nutzungsrechte gegen diese Form der Weitergabe. Legal können Informationen und Daten weiter gegeben werden, wenn diese in einer freien Lizenz veröffentlicht wurden oder eine Weitergabe ausdrücklich erwünscht ist (Freeware, Public Domain).

Anbieter hierbei ist der Tauschbörsennutzer, was zur Folge hat, dass Dateien nicht auf einem zentralen Server gespeichert sind, sondern dezentral auf die Benutzer verteilt sind (Peer-to-Peer). Die meisten Tauschbörsen arbeiten nach dem Client-Server-Prinzip, was bedeutet, dass ein Indexserver die einzelnen Dateien und ihre Anbieter genau lokalisieren kann. Bestimmte Tauschbörsen versuchen jedoch mit Hilfe bestimmter Techniken dies zu verhindern, sodass der Anbieter einer Datei nicht bestimmt werden kann.

Verschiedene Arten von Tauschbörsen

Begonnen haben Tauschbörsen im Sinne von computergestützter Verbreitung von Information und kreativen Werken mit zentral organisierten Netzwerken (z.B. Napster). Seit geraumer Zeit gibt es jedoch peer-to-peer-Netzwerke die ohne zentralen Server funktionieren (z.B. Kademlia). Hier ist prinzipiell jeder Teilnehmer Client und Server (Nutzer und Anbieter) zugleich. Damit wird eine völlige Dezentralisierung des Netzwerkes erreicht. Beispiele für diese Technik sind: Limewire, Gnutella, GNUnet ...

Napster war die erste populäre Tauschbörse. Millionen von Benutzern tauschten Musik, bis im Jahre 2000 einige Musikbands Klage gegen Napster einreichten und die Tauschbörse geschlossen wurde. Versuche Napster in eine kostenpflichtige, legale Tauschbörse umzuwandeln, schlugen fehl, da nur wenige Plattenfirmen bereit waren, ihre Musik zu lizenzieren. Die heutigen Nachfolge-Netzwerke wie Gnutella, eDonkey und FastTrack (Kazaa etc.) weisen mittlerweile jedoch deutlich mehr Nutzer auf, als Napster jemals hatte.

Darüber hinaus gibt es auch Netzwerke die nicht nur versuchen *dezentralisiert* zu arbeiten und dadurch von kontrollierenden Institutionen weitgehend unabhängig zu sein, sondern auch versuchen *Anonymität* ihrer Teilnehmer und Kontrolle der *Authentizität* des angebotenen Inhaltes zu bieten. (z.B. GNUnet und Freenet)

Tauschbörsensoftware ist legal

In den Niederlanden ist die Software für die umstrittene Tauschbörse KaZaA im Dezember 2003 für legal erklärt worden (bzw. der Anbieter der KaZaA-Software kann laut diesem Urteil nicht für die Urheberrechts-Verletzungen der Software-Nutzer verantwortlich gemacht werden). Der Hoge Raad, der höchste Gerichtshof des Landes, hat es abgelehnt, eine Klage der niederländischen Verwertungsgesellschaft für Wort und Ton, Buma/Stemra, gegen die beiden KaZaA-Gründer neu zu verhandeln. Das bedeutet jedoch nur dass die Software an sich nicht illegal ist und ihr Autor nicht für Dinge haftbar gemacht werden darf die mit seiner Software ermöglicht werden, **nicht** dass jeglich Benutzung der Software legal ist.

Weitere Arten von Tauschbörsen

Neben den populären Tauschbörsen für Dateien gibt es im Internet auch Tauschbörsen für den traditionellen Tausch von Original-Medien. So ermöglicht CdZirkel seit 2001 einen automatisierten (Ring-) Tausch von Musik-CDs, DVDs und Büchern. Noch einen Schritt weiter geht Bookcrossing das dazu ermuntert seine Bücher ganz freizugeben.

Bekannte Tauschbörsensoftware

- BitTorrent : http://bitconjurer.org/BitTorrent/
- eDonkey: http://www.edonkey2000.com

- eMule: http://www.emule-project.net
- Gnutella: http://www.gnutelliums.com
- GNUnet: http://www.ovmj.org/GNUnet/
- Kazaa: http://www.kazaa.com
- Napster: http://www.napster.com

NAPSTER

Die Musiktauschbörse **Napster** wurde 1998 von Shawn Fanning programmiert, um leichter über das Internet MP3-Musikdateien austauschen zu können. Revolutionär war dabei sein Peer-to-Peer-Ansatz. Die Napster-Software durchsuchte den Rechner, auf dem sie installiert war, nach MP3-Dateien und meldete die Ergebnisse an einen zentralen Server im Internet, wo auch die Angebote und Suchanfragen der anderen Teilnehmer eingingen. Der Server meldete als Ergebnis auf eine Anfrage die IP-Adressen der Computer zurück, die die gesuchte Musikdatei anboten. Die beiden Clients konnten sich daraufhin direkt miteinander verbinden (Peer-to-Peer) und das Musikstück übermitteln. Ein multiples Laden von mehreren Quellen, wie es später bei anderen Musiktauschbörsen eingeführt wurde, war aber nicht möglich.

Zeitweilig war Napster die am schnellsten wachsende Community (Gemeinschaft) des Internets. Dies erklärt sich durch die Sicherheit und scheinbare Kostenfreiheit des Systems sowie natürlich die attraktiven Inhalte. Kurz vor ihrem Ableben umfasste die Napster-Community circa 38 Millionen Nutzer weltweit.

Als verhängnisvoll für Napster erwies sich das Server-Client-System, das auf zentrale Rechner zur Weitervermittlung der Suchanfragen angewiesen war. Dadurch konnten die Rechteverwerter der Musikindustrie, sowie die RIAA (Recording Industry Association of America) Napster mit Klagen überziehen und die Stilllegung der Server verlangen. Schließlich wurde Napster zur Installation von Filtersoftware gezwungen, die aber nie richtig funktionierte, da die Benutzer erfindungsreich mit Dateiumbenennungen die Filter umgehen konnten (Metallica zu EtallicaM oder acillatem etc.) Zuvor hatte sich Bertelsmann in einem damals sensationellen Coup bei Napster eingekauft, zunächst als Kredit mit der Option, später den Kredit in einen Anteil umwandeln zu dürfen. Am Ende wurde Napster jedoch abgeschaltet. Mittlerweile wurde Napster unter dem Namen **Napster 2.0** als kostenpflichtiger Dienst neu gestartet.

Während der Querelen zwischen Napster und der Musikindustrie kamen alternative P2P-Programme auf, z. B. verschiedene Programme, die auf dem Gnutella-Netzwerk basieren. Das Neue an ihnen ist, dass nun keine zentralen Rechner zur Weitervermittlung der Suchanfragen mehr nötig sind.

Im Februar 2003 wurde von mehreren US-amerikanischen Komponisten und Plattenfirmen eine Klage gegen den Bertelsmann Verlag eingereicht. Dem Verlag wird vorgeworfen durch

den Kauf und die Unterstützung von Napster die Verbreitung von Raubkopien ermöglicht zu haben und damit großen wirtschaftlichen Schaden angerichtet zu haben.

Das heute kommerzielle Napster findet sich im Internet unter http://www.napster.com

BitTorrent

BitTorrent ist ein in der Programmiersprache Python geschriebenes Filesharing-Programm, das besonders für große Dateien jenseits von 100MB geeignet ist. Mittlerweile steht eine Reihe von alternativen Programmen zur Verfügung, die das *BitTorrent-Protokoll* implementiert haben.

Im Vergleich zum normalen Download einer Datei per HTTP oder FTP wird der Server des Anbieters weniger belastet und Kosten gespart: Die Upload-Kapazitäten der Herunterladenden werden mit genutzt. Dateien werden also nicht mehr nur sternförmig von einem Server verteilt, sondern auch zwischen den Nutzern verteilt (Peer-to-Peer), wodurch diese ihre Dateien schneller erhalten können.

BitTorrent besteht aus zwei Teilen: Das Server-Programm, genannt Tracker, verwaltet Informationen zu einer oder mehreren Dateien. Der herunterladenden Client erfährt vom Tracker, wer sonst noch die Datei herunterlädt und verteilt. Sobald ein Client ein kleines Stück der Datei erhalten und die Prüfsumme verifiziert hat, meldet er dies dem Tracker und kann dieses Datei-Stück schon an andere Clients weitergeben.

Um eine Datei herunterladen zu können, benötigt der Client eine Torrent-Datei (Dateiendung *.torrent*). In dieser befindet sich die Internetadresse des Trackers sowie Dateiname, Größe und Prüfsummen der herunterzuladenden Datei. Eine Torrent-Datei kann auch Informationen über mehrere Dateien beinhalten. Torrent-Dateien sind wenige Kilobytes groß und liegen üblicherweise auf der Homepage des Anbieters zum Download bereit. Löscht der Anbieter den Torrent aus dem Tracker oder geht der Kontakt zum Tracker verloren, können die Clients keinen neuen Kontakt zu anderen Clients mehr aufbauen, der Austausch zwischen schon bekannten Clients kann aber fortgeführt werden.

Im Gegensatz zu anderer Filesharing-Software kann ein Betreiber bei BitTorrent den Inhalt seines Trackers selbst bestimmen. Ein Tracker verwaltet nur die Dateien, die vom Betreiber freigeschaltet werden. Die einzelnen Tracker stehen nicht in Verbindung zueinander, es existiert daher kein gemeinsames Netz. Anbieter können sich so von fremden, möglicherweise illegalen Inhalten distanzieren.

Red Hat gehörte zu den ersten Linux-Distributionen, die auch mit BitTorrent die Distribution verteilte.

Das Programm findet sich unter http://bitconjurer.org/BitTorrent/

SUCHMASCHINE

Eine **Suchmaschine** ist ein Programm zur Recherche von Dokumenten, die in einem Computer oder einem Computernetzwerk wie z. B. dem World Wide Web gespeichert sind. Nach Eingabe einer Suchanfrage - bestehend aus Schlagwoertern und der Art ihrer logischen Verknüpfung - liefert eine S. eine Liste von Verweisen auf relevante Dokumente oder Webseiten, meist dargestellt mit Titel und einem kurzen Auszug des jeweiligen Dokuments.

Im Gegensatz zu gelegentlich vom Begriff der Suchmaschine miterfassten Such-Kataloge wie z. B. Yahoo!), die Dokumente hierarchisch in einem Inhaltsverzeichnis nach Themen organisieren, ist die Recherche mit Suchmaschinen im eigentlichen Sinne (z. B. AltaVista, Google, Fireball) nach Schlagworten bzw. Suchbegriffen organisiert.

Die zur Suche noetige Datenbasis gewinnen Suchmaschinen meist durch automatisierten Abruf und Indexierung einer großen Anzahl von Webseiten sowie Verfolgung der darin enthaltenen Verweise, d. h. Links, oder im Fall von Metasuchmaschinen durch Rueckgriff auf mehrere andere Suchmaschinen.

Die Darstellung der Suchergebnisse geschieht sortiert nach Relevanz, wofür jede Suchmaschine ihre eigenen, geheimen Kriterien heranzieht. Das koennen z. B. Häufigkeit und Stellung der Suchbegriffe im jeweiligen gefundenen Dokument, die Einstufung und Anzahl der zitierten Seiten, sowie die Haeufigkeit von Verweisen anderer Dokumente auf das im Suchergebnis enthaltene Dokument sein.

Manche Suchmaschinen sortieren Suchergebnisse nicht nur nach Relevanz für die Suchanfrage sondern lassen gegen Bezahlung auch Einflussnahme auf ihre Ausgabe zu.

Eine der ersten Suchmaschinen war Lycos, die 1994 als Universitätsprojekt gestartet ist. Vorher gab es schon andere, z. B. *Archie* oder *Veronica*, die aber nicht mit einer graphischen Web-Oberfläche ausgestattet waren.

Herausforderungen für Suchmaschinen:

- **Mehrdeutigkeit:** Suchanfragen von Benutzern und Inhalte von Webseiten sind häufig unklar und mehrdeutig. Die Suchmaschine muss natürliche Texte zu einem gewissen Grad "verstehen", um relevante Antworten liefern zu können. Idealerweise wird eine Suchmaschine zur Anfrage "Rechner Linux" auch Seiten liefern, die das Wort "Rechner" gar nicht enthalten und statt dessen "Computer" benutzen.
- Datenmenge: Das Web wächst schneller als die Suchmaschinen mit der derzeitigen Technik indizieren können.
- Aktualität: Viele Webseiten werden häufig aktualisiert, was die Suchmaschinen zwingt, diese Seiten immer wieder zu besuchen.
- Spam: Mittels Suchmaschinen-Spamming, d. h. mit unlauteren Methoden (z. B. versteckten Texten und unsichtbaren Links auf der Webseite, automatischen Weiterleitungen von gut platzierten Brückenseiten auf unerwünschte Seiten), versuchen selbsternannte "Suchmaschinen-Optimierer", den Algorithmus einer Suchmaschine, der

für die automatische Bewertung und günstige Platzierung von Webseiten in den Ergebnislisten zuständig ist, auszutricksen. Mit anderen Worten: Die Manipulateure versuchen, ihre eigenen Internetseiten (oder die ihrer Kundschaft), die normalerweise nicht unter den besten Ergebnissen einer Suchabfrage angezeigt würden, durch Anwendung diverser Spamming-Methoden in den Ergebnislisten der Suchmaschinen möglichst weit vorn zu platzieren. Dadurch liefert die Suchmaschine auf ihren ersten Ergebnisseiten keine relevanten, vom Surfer gewünschten Resultate mehr, sondern präsentiert dem Internetnutzer die Seiten des Spammers. Die meisten Suchmaschinen versuchen, sich gegen diese Form der Manipulation zu wehren.

World Wide Web

Das **World Wide Web** (kurz das **Web** oder **WWW**, wörtlich: *Weltweites Gewebe/Netz*) ist ein Hypertext-System, das über das Internet abgerufen werden kann.

Hierzu benötigt man einen Webbrowser, um die Daten vom Webserver zu holen und z. B. auf dem Bildschirm anzuzeigen. Der Benutzer kann den Hyperlinks im Dokument folgen, die auf andere Dokumente verweisen, egal ob sie auf dem gleichen Webserver oder einen anderen gespeichert sind. Hierdurch ergibt sich ein *weltweites Netz oder Gewebe aus Webseiten*. Das Verfolgen der Hyperlinks wird oft als „Surfen im Web“ bezeichnet.

Das WWW wird im allgemeinen Sprachgebrauch oft mit dem Internet gleichgesetzt, obwohl es nur eine mögliche Nutzung des Internets ist sowie im Gegenzug das Internet nur eine von verschiedenen möglichen Serververbünden. Es gibt durchaus Internet-Dienste, die nicht in das WWW integriert sind (z. B. IRC und Telnet). Dazu beigetragen haben nicht zuletzt die Webbrowser, die nicht nur das eigentliche HTTP-Protokoll (siehe unten) benutzen können, sondern auch noch andere Dienste wie FTP dem Nutzer zugänglich machen.

Geschichte

Das Web entstand 1989 als Projekt am CERN, an dem Tim Berners-Lee ein Hypertext-System aufbaute. Das ursprüngliche Ziel des Systems war es, Forschungsergebnisse auf einfache Art und Weise mit Kollegen auszutauschen. Das dem Hypertext zugrunde liegende Konzept stammt von früheren Entwicklungen ab, wie Ted Nelsons Project Xanadu, Vannevar Bushs „memex“ Machinenidee und dem Note Code Project. Das World Wide Web besitzt klare Unterschiede im Vergleich zu damaligen Hypertext-Systemen (Note Code benutzte bspw. eine einfache und lesbare Syntax und sogar semantische Deskriptoren). Das WWW benötigt nur unidirektionale links, anstatt bidirektionale, was es jedem ermöglicht einen Link auf eine Resource zu setzen, ohne daß der Besitzer dieser eingreifen muß. Zudem, anders als andere Protokolle, wie Hypercard oder Gopher, baut das World Wide Web auf einem freien Protokoll auf, was die Entwicklung von Servern und Clients ohne Lizenzierungsbeschränkungen möglich machte.

Das erste WWW-Programm am CERN konnte nur Text anzeigen, aber spätere Browser wie Pei Weis Viola (1992) fügten die Fähigkeit Grafiken anzuzeigen zu. Marc Andreessen vom

NCSA veröffentlichte einen Browser namens „Mosaic für X“ 1993, der dem Internet eine riesige Popularität und ein riesiges Wachstum bescherte. Marc Andreesen gründete die Firma „Mosaic Communications Corporation“, später „Netscape Communication“. Zusätzliche Merkmale wie dynamischer Inhalt, Musik und Animation können nun in modernen Browsern angetroffen werden.

Funktionsweise

Das WWW basiert auf drei Kernstandards:

1. HTTP als Protokoll, mit dem der Browser Informationen vom Webserver anfordern kann,
2. HTML als Dokumentbeschreibungssprache, die festlegt, wie die Information gegliedert ist und wie die Dokumente verknüpft sind (Hyperlinks), und
3. URIs als eindeutige Adresse bzw. Bezeichnung einer Ressource (z. B. einer Webseite), die in Hyperlinks Verwendung findet.

Folgende Standards kamen später hinzu:

1. Cascading Style Sheets (CSS) legen das Aussehen der Elemente einer Webseite fest, wobei Darstellung und Inhalt getrennt werden.
2. ECMAScript (auch Javascript oder JScript) ist eine Programmiersprache mit Anweisungen für den Browser und erlaubt das Einbetten von Programmen (Skripte). Dadurch können Webseiten mit Hilfe des Document Object Models (DOM) dynamisch geändert werden. Skripte sind üblicherweise kleine Programmschnipsel, können aber auch als Client Manager mit Hilfe des DOM die vollständige Kontrolle über die Anzeige übernehmen.

Das World Wide Web Consortium, das von Tim Berners-Lee geleitet wird, entwickelt und verfechtet hierbei den HTML- und CSS-Standard; andere Standards stammen von der Internet Engineering Task Force, der ECMA oder Herstellern wie Sun Microsystems.

Das WWW wurde und wird durch andere Technologien ergänzt. Schon sehr früh wurden Bilder zur Illustration benutzt; die vorherrschenden Formate sind GIF, PNG und JPEG.

Außerdem erlaubt HTML das Einbetten oder Verlinken nahezu aller Dateitypen, die der Browser mittels Ergänzungsmodulen darstellen kann. Hierdurch lassen sich Multimediainhalte von Animationen bis hin zu Musik und Videos darstellen. Ferner erlaubt Java das Einbetten von Programmen, die auf dem Computer des WWW-Benutzers ablaufen.

Weitere beliebte Formate sind PDF oder Flash.

Dynamische Webseiten und Webanwendungen

Mit Hilfe der dynamischen WWW-Seiten, die entweder auf dem Webserver (mittels CGI oder Skriptsprachen wie PHP oder Perl) oder durch den Browser (mittels ECMAScript) generiert werden, kann das WWW als Oberfläche für verteilte Programme angesehen

dienen: Ein Programm wird nicht mehr konventionell lokal auf dem Rechner gestartet, sondern ist einfach eine Menge von dynamischen WWW-Seiten, die mittels eines Web Browsers betrachtet und bedient werden können. Vorteilhaft ist hier, dass die Programme nicht mehr auf den einzelnen Rechnern verteilt und dort (dezentral) administriert werden müssen.

Nachteilig sind die begrenzten Ausdrucksmöglichkeiten von WWW-Seiten, so dass die Bedienbarkeit von Programmen in Form von Internetseiten im Allgemeinen nicht an die von konventionellen Programmen heranreicht.

Ferner ist eine Entwicklung zu beobachten, die dazu führt, dass immer mehr Dienste, die ursprünglich vom WWW getrennt waren und mit getrennten Programmen genutzt werden, immer öfter über die technischen Möglichkeiten des WWW angeboten werden und mittels eines Browsers genutzt werden können:

So wird E-Mail oft als Webmail genutzt; Webforen ersetzen das Usenet und Webchats das IRC.

Kompatibilität und Zugänglichkeit

Oft führten Browser-Hersteller neue Möglichkeiten ein, ohne auf eine Standardisierung zu warten. Umgekehrt werden jedoch immer noch nicht alle Teile von Standards wie HTML oder CSS (richtig) implementiert. Dies führt zu Inkompatibilitäten zwischen bestimmten Webseiten und manchen Browsern.

Außerdem ging durch die unzähligen Ad-Hoc-Erweiterungen von HTML der Vorteil dieser Sprache verloren, der darin besteht, dass Inhalt und Darstellung getrennt werden. Dadurch können die in HTML ausgezeichneten Inhalte optimal für das Ausgabegerät - ob Bildschirm, Display des Mobiltelefons oder Sprachausgabe (für Benutzer mit Sehschwierigkeiten) - aufbereitet werden.

Das W3C und andere Initiativen forcieren daher eine Entwicklung in die Richtung XHTML/XML und CSS, um diese Vorteile von HTML wieder zu erlangen.

Weblinks

- Zeitleiste der Geschichte des WWW (*http://www.dejavu.org/*) mit Emulationen alter Dienste und Browser
- Das Projekt WWW (*http://groups.google.com/groups?selm=6487%40cernvax.cern.ch*) und die erste Vorstellung im Usenet (1991, englisch)

Webseite

Webseiten sind Dateien, die mit einem Browser angezeigt werden können. Ihr wesentliches Merkmal ist, dass sie Verweise (Hyperlinks) auf andere Webseiten enthalten und somit einen Hypertext darstellen. Die Gesamtheit der Webseiten unter einer Webadresse wird als

Website oder auch *Webpräsenz* bezeichnet. Webseite und Website sind also *keine Synonyme*. Jede Webseite kann über eine eindeutige URL aufgerufen werden. Webseiten können statischen (festen) Inhalt oder dynamischen Inhalt haben, der bei jedem Aufruf neu generiert wird.

Webseiten bestehen zumeist aus strukturiertem Text. Bilder und Multimediaelemente können ebenfalls in ihnen dargestellt werden. Webseiten werden aus dem Internet oder einem Intranet geladen, können aber auch auf einem lokalen Speichermedium, zum Beispiel einer Festplatte abgelegt sein.

HTML und XML werden genutzt, um Webseiten zu schreiben; CSS um sie zu gestalten. Diese **Trennung von Inhalt und Gestaltung** wird jedoch häufig missachtet. Dies führt dazu, dass Webseiten von den verschiedenen Browsern nicht benutzergerecht dargestellt werden. Im Extremfall wird eine Webseite von einem oder mehreren Browsern überhaupt nicht angezeigt.

Um Webseiten zu finden, werden Suchmaschinen genutzt.

Hypertext

Als **Hypertext** bezeichnet man die *nicht-lineare* Organisation von heterogenen Objekten, deren netzartige Struktur durch logische Verbindungen (= Verweise, Links) zwischen atomisierten Wissenseinheiten (= Knoten, z. B. Texte oder Textteile) hergestellt wird (*Verweis-Knoten-Konzept*).

Die Begriffe *Hypert*ext und Hypermedia werden meistens synonym benutzt; *Hypertext* betont dabei jedoch den textuellen Anteil, *Hypermedia* dagegen mehr den multimedialen.

Die Wikipedia ist ein Beispiel für einen komplexen Hypertext.

Entwicklung und Geschichte

Das erste Mal wurde dieses Konzept wohl 1945 von Vannevar Bush in einem Artikel „*As We May Think*“ im Journal The Atlantic Monthly erwähnt. Er sprach darin über ein zukünftiges System Memex (für Memory Extender), das das Wissen eines bestimmten Gebietes elektronisch aufbereitet leicht zugänglich darstellen kann. Der Kernpunkt ist, dass man mit elektronischer Hilfe den Verweisen leicht folgen kann und dass Bücher und Filme aus einer Bibliothek angezeigt werden.

Der Gesellschaftswissenschaftler Ted Nelson prägte den Begriff „*Hypertext*“ im Jahr 1965.

Eines der ersten Hypertextsysteme, das einer größeren Gruppe zugänglich war, war HyperCard der Firma Apple Computer, das mit den Apple Macintosh Computern ausgeliefert wurde.

Das heute am weitesten verbreitete System ist das World Wide Web, obwohl ihm einige wichtige Funktionen früherer Hypertextsysteme fehlen. So ist zum Beispiel das Problem der

„toten“ Links (Dead Links, Dangling Links) im WWW ungelöst. Auch die Implementierung der Uniform Resource Identifiers (URIs) ist über die im Web gebräuchlichen URLs nur unvollständig erfüllt. Im Gegenzug erlaubt das WWW aber auch das Einbinden von nichtsprachlicher Information. Dadurch ist es, obwohl auf Hypertext beruhend, streng genommen ein Hypermedia-System.
Die Sprache, in der die Texte des World Wide Web beschrieben werden, heißt HTML. Auch viele Software-Dokumentationen (Hilfe-Texte) sind als Hypertext konzipiert.

HYPERLINK

Ein **Hyperlink** (in der Regel kurz **Link**, aus dem Englischen für Verknüpfung, Verbindung, Verweis) ist ein Verweis von einem Webdokumenten oder Teilen davon durch eine entsprechende Markierung auf ein anderes Webdokument (typischerweise einer HTML-Seite). Auch in anderen Hypertext-Medien gibt es Hyperlinks.

Diese Ressource kann z. B. eine Datei (Webseite, Bild, Audiodatei, Videodatei, etc.) oder eine dynamisch beim Aufruf erstellte HTML-Seite sein. Es sind aber auch Verweise auf Dienste möglich (FTP, whois, gopher...).

Die Interpretation der Links übernimmt ein (Web-)Browser, wie z.B. der Internet Explorer von Microsoft, der Netscape Navigator oder Opera.

Verlinken ist ein denglisches Wort, abgeleitet von Englisch link (Verbindung, Kettenglied). Es bezeichnet das Erstellen von **Hyperlink**s zwischen verschiedenen Dokumenten; in der Regel handelt es sich um HTML-Dokumente, die über HTTP (das Internet) erreichbar sind. Verlinken kann man aber beispielsweise auch PDF- oder Info-Dokumente (Texinfo) und als Übermittlungsprotokolle können z.B. auch HTTPS oder FTP in Frage kommen.

Auch die Seiten der Wikipedia sind mittels modifizierten WikiLinks ([[...]]-Syntax) verlinkt.

In der Regel kann man große, verlinkte Netze als gerichtete zyklischen Graphen repräsententieren, in denen die Kanten von den Hyperlinks und die Ecken oder Knoten von den referenzierten Seiten gebildet werden.

WEBLINK

- http://www.linksandlaw.com/linkingcases-linkingpolicies-beispiele.htm - Nutzungsbedingungen auf deutschsprachigen Webseiten, die das Setzen von Hyperlinks einschränken

UNIFORM RESOURCE IDENTIFIER

Ein **Uniform Resource Identifier** (**URI**), engl.: einheitlicher Bezeichner für Ressourcen) ist eine Zeichenfolge, die zur Identifizierung einer abstrakten oder physikalischen Ressource

dient. URIs werden zur Bezeichnung von Ressourcen (wie Webseiten oder anderen Dateien, aber auch z. B. E-Mail-Empfängern) im WWW eingesetzt.

URIs können als Zeichenfolge (kodiert mit einem Zeichensatz) in digitale Dokumente, insbesondere solche im HTML-Format eingebunden oder auch von Hand auf Papier aufgeschrieben werden. Ein Verweis von einer Webseite auf eine andere nennt man Link.

Derzeit können URIs nur aus druckbaren ASCII-Zeichen bestehen; eine Erweiterung zu so genannten *Internationalized Resource Identifiers* (IRIs) ist jedoch geplant.

Aufbau

Der erste Teil eines URIs (vor dem Doppelpunkt) gibt den Typ des URIs an, der die Interpretation des folgenden Teils festlegt:

```
<Schema>:<Schemen-spezifischer Teil>
```

Viele URI-Schemata wie `http` oder `ftp` besitzen einen hierarchischen Aufbau:

```
<Schema>://<Server>:<Port>/<Pfad,...>?<Anfrage>
```

<*Server*> gibt hierbei bei Schemata, die ein TCP- oder UDP-basiertes Protokoll verwenden, den Domainnamen des Servers an; <*Port*> den TCP-Port. Das bedeutendste Schema ist `http` für das Hypertext Transfer Protocol.

Hierarchische URIs können ferner relativ zu einem Basis-URI angegeben werden. Hierbei werden Schema, Server und Port sowie gegebenenfalls Teile des Pfades weggelassen.

An URIs kann, abgetrennt durch `#`, auch ein Fragmentbezeichner angehängt werden. Eine Kombination aus URI und Fragmentbezeichner wird als **URI-Referenz** bezeichnet.

Schemata

Unter anderem sind folgende Schemata definiert:

- `http` - Hypertext Transfer Protocol
- `ftp` - File Transfer Protocol
- `mailto` - E-Mail-Adresse
- `urn` - Uniform Resource Names (URNs)
- `tel` - Telefonnummer
- `news` - Newsgroup oder Newsartikel
- `data` - direkt eingebettete Daten
- Vollständige Liste: http://www.iana.org/assignments/uri-schemes (englisch)

URIs, URLs und URNs

Man unterscheidet folgende Unterarten von URIs:

- **Uniform Resource Locators** (**URLs**) identifizieren eine Ressource über ihren primären Zugriffsmechanismus, geben also den *Ort* (engl. *location*) der Ressource im Netz an. Beispiele hierfür sind `http` oder `ftp`. URLs waren ursprünglich die einzige Art von URIs, weshalb der Begriff URL oft als gleichbedeutend zu URI verwendet wird.
- **Uniform Resource Names** (**URNs**) mit dem URI-Schema `urn` identifizieren eine Ressource mittels eines vorhanden oder frei zu vergebenden Namens, z. B. `urn:isbn`.

Ursprünglich sollte jeder URI in eine dieser beiden Klassen (oder weitere noch zu definierende) eingeteilt werden. Diese strenge Aufteilung wurde jedoch aufgegeben, da sie unnötig ist und einige Schemata (wie `data`) in keine der beiden Klassen passen.

Manche Schemata (wie `mailto`), die früher als URL bezeichnet wurden, sind heute keiner der beiden Klassen zuzuordnen.

Webbrowser

Webbrowser oder **Browser** sind Computerprogramme, die zum Betrachten verschiedener (v. a. textorientierter) Inhalte benutzt werden. Wegen der Metapher, dass man rasend schnell durch das Web braust und der Wortähnlichkeit zu Browser, wird manchmal ironisierend der Begriff **Brauser** verwendet. Vorwiegend werden sie verwendet, um Websites im World Wide Web anzuzeigen.

Ursprünglich bezeichnet der aus dem Englischen entlehnte Begriff *browsen* (en: *to browse* = stöbern, sich umsehen, schmökern) am Computer lediglich das Benutzen von Navigations-Elementen (Vor, Zurück, Index, ...) zum Lesen von Texten. Erweitert wurde dieser Begriff später durch das Aufkommen von Hypertext, bei dem man bestimmte Worte auswählen kann (Link), um zu einem anderen Text zu gelangen. Später kamen dann Funktionen zum Bildbetrachten dazu und auch sogenannte Imagemaps, bei denen man auf einer Computergrafik einen Bereich (z. B. bei einer Weltkarte) klickt und dadurch zu einer Seite über ein bestimmtes Land gelangt.

Mit dem Trend zu Multimedia wurde der Browser eine der zentralen Schnittstellen auf einem PC. Man kann verschiedene Dateien abspielen lassen wie Musik oder Radio. Insbesondere die Verbreitung von Breitband-Internetzugängen förderte diese zentrale Funktion heutiger Browser. Somit verschwimmt zunehmend auch der Unterschied zu einem Dateimanager, der ursprünglich ausschließlich zum Öffnen, Kopieren oder Löschen von Dateien verwendet wurde. Viele Dateimanager beinhalten heute auch Browser-Funktionen (*Datei-Browser*) bzw. können auch zum Anzeigen von Dokumenten verwendet werden.

Neben der Erweiterung unterstützter Dateiformate, wurden auch die Funktionen erweitert. So unterstützen viele Browser neben HTTP auch u. a. die Protokolle FTP und Gopher. Somit kann man mit Browser heutzutage auch Programme oder Dateien herunterladen.

Einige Browser beinhalten auch Funktionen für E-Mail oder Usenet. Andere Browser liefern für diese Funktionen externe Programme.

Browser-Arten

Man unterscheidet zwischen textbasierten und grafischen Browsern.

Textbasierte Browser können nur einfachen Text und Textformate wie HTML oder XML interpretieren und darstellen. Meist bieten sie die Möglichkeit Computergrafik-Dokumente abzuspeichern oder mit externen Programmen darzustellen. Textbrowser sind besonders für Blinde nützlich, die das Internet barrierefrei nutzen wollen. Mit ihnen kann, etwa für Blinde, Text leicht in Sprache umgewandelt werden. Sie eignen sich besonders zu schnellen Recherche, da Bilder, Werbung u.ä. gar nicht geladen werden.

Beispiele für textbasierte Browser sind Lynx, Links und w3m

Grafische Browser verlangen als Basis auch ein Betriebssystem mit einem GUI (z. B. Unix mit X11, Macintosh, Windows) und werden am häufigsten verwendet. Sie zeigen Inhalte wie Computergrafiken, Filme oder Java-Applets überwiegend selbst an oder sie benutzen dazu sogenannte Plugins. Viele dieser Browser (wie Mozilla, Internet Explorer) werden heute als Browser-Suite mit integrierten Funktionen für z. B. E-Mail und Usenet ausgeliefert.

Der marktbeherrschende Browser ist momentan der von Microsoft produzierte Internet Explorer. Dieser Browser weist durch die starke Integration in das Betriebssystem Windows allerdings auch mit Abstand die meisten Sicherheitslücken auf (bzw. sind die meisten bekannt). Die größte Konkurrenz zu diesem Browser sind derzeit Mozilla bzw. Netscape, Opera, Konqueror und Safari (Browser).

Mozilla

Mozilla ist eine aus Webbrowser, E-Mail-Client und weiteren Werkzeugen bestehende Programmsammlung, die von einer Entwicklergruppe um *mozilla.org* unter der *Mozilla Public License* und der *Netscape Public License* entwickelt wird. Mozilla basiert auf dem 1998 von der Netscape Corp. freigegebenen Quellcode des Netscape Navigators; von dieser Codebasis enthält Mozilla allerdings nicht mehr viel. Mozillas (neugeschriebene) interne Layout-Maschine heißt Gecko.

Der Name *Mozilla* ist seit den Anfängen von Netscape die interne Bezeichnung für den Browser. Zum genauen Ursprung dieses Wortspiels, das sich auf den den ersten populären Webbrowser Mosaic bezieht, gibt es mehrere Erklärungen. Eine besagt, dass sich der Name aus *Mosaic Killa* zusammen setzt, laut der anderen aus *Mosaic* und *Godzilla*. Das offizielle

Netscape-Maskottchen, eine grüne Eidechse, und der bei Mozilla verwendete Dinosaurier deuten auf letztere Erklärung.

Mozilla läuft zur Zeit auf vielen verschiedenen Betriebssystemen (*Plattformen*). Dies wird dadurch ermöglicht, dass große Teile von Mozilla *plattformunabhängig* sind. Die Hauptarbeit bei der Portierung entfällt auf die **Netscape Portable Runtime**, die eine allgemeine Schnittstelle für z. B. Dateizugriff und Speicherverwaltung definiert.

Die Programmoberfläche wird in der eigens entwickelten, ebenfalls plattformunabhängigen Sprache XUL, die auf XML basiert, geschrieben. Die Elemente dieser Oberfläche werden, ähnlich wie Webseiten, durch *Gecko* dargestellt und nicht vom Betriebssystem. Dies hat den Vorteil, dass ein Entwickler für Änderungen an der Oberfläche keine Kenntnisse mehrerer Betriebssysteme benötigt. XUL ermöglicht auch, in einfacher Weise durch sogenannte *Extensions* die Mozilla-Oberfläche anzupassen oder zu erweitern. XUL unterstützt auch sogenannte Themes, die auf Basis von Webstandards wie Document Object Model und Cascading Style Sheets erstellt werden und mit denen das Aussehen von Mozilla dem eigenen Geschmack angepasst werden kann. Ein Nachteil ist jedoch, dass eine XUL-basierte Oberfläche auf älteren Rechnern langsamer läuft.

Im Gegensatz zur kommerziellen Variante "Netscape" liefert Mozilla kaum Erweiterungen, sogenannte Plugins mit, weil diese oft nicht in einer freien Lizenz verfügbar sind. Diese können jedoch nachträglich installiert werden.

Mozilla ist auch in der Lage Stylesheets auf HTML- und XML-Seiten anzuwenden. Dazu werden derzeit die Standards CSS 1 (vollständig) und CSS 2 (nahezu vollständig) und XSLT unterstützt. Daneben gibt es einen so genannten Pop-Up-Blocker und die Funktion des tab-basierten Browsens (mit Reitern als Alternative zu Fenstern).

Neben dem Browser beinhaltet Mozilla auch ein E-Mail-Modul inklusive Adressbuch. Eines der wichtigsten Features ist der seit Version 1.3 enthaltene adaptive Spam-Filter auf Basis der Bayes'schen Wahrscheinlichkeitstheorie. Nach einer Lernphase, in welcher der Benutzer Mails als "Junk" (also Spam) und "Nicht Junk" klassifiziert, weist das Programm sehr gute Trefferquoten auf.

Außerdem enthalten ist das Modul *Composer* zum Bearbeiten und Erstellen von HTML-Seiten sowie *Chatzilla* zum Chatten in IRC-Netzwerken.

Varianten und Ableitungen

Netscape Navigator Version 6.x und 7.x basieren auf Mozilla-Code. Auch Konqueror kann auf die Mozilla-Engine umgeschaltet werden.

Da der Code von Mozilla oft als zu groß und unübersichtlich kritisiert wurde bildeten sich verschiedene Projekte, die auf Gecko basierend einen schnelleren, schlankeren und leichter zu wartenden Browser kreieren wollen. Diese sind unter Anderem:

- Mozilla Firefox (früher: Phoenix, Firebird)
- Galeon (Standardbrowser von Gnome 2.2)

- Epiphany (Standardbrowser von Gnome 2.4 und 2.6)
- Camino (früher: Chimera, für Mac OS X)
- K-Meleon (für Windows)
- Skipstone

Die Mozilla Roadmap sieht vor, dass sich das Projekt mehr auf Einzelprodukte konzentriert. Dazu werden die Entwicklung des Browsers Mozilla Firefox und von Mozilla Thunderbird, dem Mail-Programm, stärker vorangetrieben. Der Hauptgrund für diese Änderung ist die bereits erwähnte Größe des Mozilla Paketes. Zudem hat aber auch die Erfahrung mit diesen beiden Projekten gezeigt, dass sich dadurch benutzerfreundlichere Programme ergeben, da sich die Entwickler auf ein einzelnes Produkt und nicht auf die ganze Suite konzentrieren.

Entwicklungsstand

Mozilla 1.0 erschien Mitte des Jahres 2002. Die derzeit aktuelle, stabile Version 1.6 erschien am 15. Januar 2004.

Am 10. März 2004 warnte das CERT der Universität Stuttgart vor Schwachstellen im Mozillacode (http://cert.uni-stuttgart.de/ticker/article-print.php?mid=1183). Kritisiert wurde vor allem, dass Schwachstellen vor allem stillschweigend behoben werden. Durch diese fehlenden Informationen nutzen Anwender in gutem Glauben ältere Versionen, die noch Mängel enthalten. Es wurde empfohlen, immer die letzte Version von Mozilla zu nutzen.

Bookmark / Lesezeichen

Im Internet versteht man unter **Lesezeichen** eine Linksammlung in Anwenderprogrammen, meist Webbrowsern.

Je nach Anwenderprogramm werden unterschiedliche Bezeichnungen verwendet, meist jedoch *Lesezeichen*, im Internet Explorer jedoch *Favoriten*. Bei englischsprachigen Programmen findet sich häufig die Bezeichnung *Bookmark*.

Lesezeichen helfen eine einmal gefundene Seite auch ohne Suchmaschine wiederzufinden. Erfahrenen Internetanwender sammeln Lesezeichen und kommen so häufig auf tausende von Einträgen.

Viele Webbrowser, darunter Mozilla und Lynx, verwalten Lesezeichen in einer HTML-Datei (meistens kein reines HTML, sondern mit speziellen Erweiterungen). Opera verwendet eine speziell formatierte Textdatei. Bei den Favoriten des Internet Explorer wird pro Eintrag eine kleine Datei real im Dateisystem in Ordnern abgelegt.

WEBDESIGN

Webdesign (auch *Webgestaltung*) umfasst die Gestaltung, den Aufbau und die Nutzerführung von Webseiten für das WWW. Dabei werden i. d. R. strukturierte Texte, die in (X) HTML verfasst sind, samt ergänzenden Grafiken mit CSS formatiert.

Der Webgestalter hat dabei die Aufgabe den besten Kompromiss zwischen den Wünschen des Auftraggebers, den Ansprüchen der Besucher und den technischen Möglichkeiten zu finden.

Dynamischer Inhalt lässt sich entweder durch serverseitige Skriptsprachen wie PHP oder Perl erzeugen oder auch durch clientseitige Erweiterungen wie Flash, Java oder JavaScript. Es besteht auch die Möglichkeit, client- und serverseitige Technologien zu kombinieren, beispielsweise PHP und Flash, um die Vorteile beider nutzen zu können. Dabei sollte man darauf achten, mit clientseitigen Erweiterungen sparsam umzugehen, da oft die notwendigen Plugins beim Benutzer nicht vorhanden sind oder JavaScript aus Sicherheitsgründen abgeschaltet wurde.

Neben der reinen Optik geht es bei der Gestaltung von Webseiten vor allem um *Nutzerfreundlichkeit*. Navigation und Aufbau der Webseiten sollen möglichst vielen Menschen entgegen kommen. Hier erfahren aber z. B. viele behinderte Menschen Nachteile, sie benötigen Webseiten, die barrierefrei gestaltet sind.

Zur Nutzerfreundlichkeit kommt die Zugänglichkeit (Accessibility), die sich in der Vermeidung von Techniken manifestiert, durch die Informationen nur mit einem bestimmten Browser erreichbar sind, oder durch das Schaffen von (Text-)Alternativen zu multimedialen Inhalten. Flash und andere Browsererweiterungen müssen deswegen nicht grundsätzlich vermieden werden, es sollte aber sichergestellt sein, dass der Inhalt auch ohne diese Techniken voll abrufbar bleibt.

> Ohne *Accessibility* keine *Usability*!

Ein wichtiger Aspekt beim Webdesign ist eine korrekte Textauszeichnung und Kenntnisse in Webtypografie. Während Webseiten für die Browsergenerationen 4 (Netscape 4 und Internet Explorer 4) noch sehr unterschiedlich programmiert wurden, kann der Webentwickler in den aktuellen Versionen (Mozilla, Internet Explorer 6, Opera, Konqueror, usw.) sich voll und ganz auf die Standards des W3C verlassen werden. Noch gibt es ein paar Ausnahmen, aber die dürften bald beseitigt sein.

Webdesign unterscheidet sich vom Design für andere Medien vor allem durch diese starke Leserorientierung. Sie ist darin begründet, dass Websurfer meist gezielt nach Informationen suchen und eine Site nur so lange besuchen wie unbedingt nötig.

AUSZEICHNUNGSSPRACHE

Eine **Auszeichnungssprache** (auch *Markup Languange* (*ML*)) dient zur Beschreibung von Informationen oder des Verfahrens oder der Schritte, die zur Darstellung nötig sind.

Ursprünglich diente das Auszeichnen von Texten den Setzern aus der Druckindustrie als Anweisung.

Es gibt zwei Gruppen von Markup Languages. Die eine Gruppe (Descriptive Markup Language, *kurz: DML*) beschreibt Informationen. Die andere Gruppe (Procedural Markup Language, *kurz: PML*) beschreibt das Verfahren oder die Schritte, die zur Darstellung nötig sind.

Mit Hilfe einer *ML* werden die Eigenschaften, Zugehörigkeiten und Verfahren von bestimmten Wörtern, Sätzen und Abschnitten eines Textes beschrieben bzw. zugeteilt. Die Quelltexte werden in der Regel im ASCII-Code oder Unicode (dann meist UTF8) verfasst. Oft werden die Sonderzeichen auch von der ML beschrieben, meist mit Hilfe einer nummerische Zuweisung (unicode) oder durch Benennung (z. B. LaTeX: \mu für µ; WikiPedia: µ für µ).

Beispiel für ...	HTML	LaTeX	WikiPedia
Überschrift	Überschrift	\section{Überschrift)	= Überschrift =
Aufzählung	 - Punkt 1 - Punkt 2 - Punkt 3 	\begin{itemize} \item Punkt 1 \item Punkt 2 \item Punkt 3 \end{itemize}	* Punkt 1 * Punkt 2 * Punkt 3
Bold	FETT	\bf{FETT}	'''FETT'''

Descriptive Markup Languages sind unter anderem die in SGML oder XML definierten Sprachen HTML, DocBook, TEI, MathML oder SVG, aber auch TeX, wenn es mit Makropaket LaTeX verwendet wird. Auch das WikiPedia-Projekt hat eine eigene ML. Diese Formate erleichtern die Erstellung von Formatierungen oder auch Tabellen für die ansonsten Kenntnisse in HTML nötig wären.

Procedual Markup Languages sind unter anderem TeX, PDF und PostScript. Diese Formate weisen, unter anderem den Drucker, an, wie das Dokument ausgedruckt werden soll. Die Auszeichnungssprachen besitzen wie jede Sprache eine Syntax, Grammatik und eine Semantik, die je nach Sprache varriert. Beispiele befinden sich in der Vergleichstabelle.

HTML

Die **Hypertext Markup Language** (HTML) ist ein Dokumentenformat zur Auszeichung von Hypertext im World Wide Web und wurde 1989 von Tim Berners-Lee am CERN in Genf festgelegt. Sie basiert dabei auf der Metasprache SGML, die zur Definition von Auszeichnungssprachen verwendet wird. HTML ist also eine Auszeichnungssprache zur Beschreibung von Informationen in Hypertexten. Auch wenn sich ältere HTML-Versionen dafür verwenden lassen, geht es in HTML also nicht um die optische Gestaltung eines Textes. Vielmehr lassen sich einzelnen Textbereichen einzelne semantische Funktionen zuordnen (z. B. Überschrift), deren optische Gestaltung hierbei nebensächlich ist und mit CSS festgelegt werden kann. Beispielsweise gibt es ein HTML-Element zur Betonung von Textbereichen. Was ein Webbrowser daraus macht, ist Sache des Webbrowsers, nicht Sache von HTML. Ob ein mit diesem Element ausgezeichneter Textabschnitt vom Webbrowser durch Kursivschrift, Fettschrift oder eine auffällige Farbe realisiert wird, kann jedoch mit CSS neu festgelegt werden. Der nachfolgende Text soll diese Eigenschaften von HTML konkretisieren. Mittlerweile wurde die letzte Version des HTML-Standards (HTML 4.01) in der Meta-Sprache XML neu formuliert. Das daraus entstandene XHTML 1.0 genügt den im Vergleich zu SGML strengeren syntaktischen Regeln von XML, ist aber in seinen drei DTD-Varianten (Strict, Transitional und Frameset, *siehe unten*) semantisch mit der jeweils entsprechenden DTD-Variante von HTML 4.01 identisch. Aktuell liegt der XHTML-Standard in der Version 1.1 vor, der eine zusätzliche Modularisierung der Elemente einführt.

Überblick

Namensgebend sind die Hypertext-Elemente, die zum Verweis auf andere Textstellen oder auf ein anderes Dokument dienen. Zur Adressierung anderer Dokumente im Internet werden innerhalb des Dokumentes Hyperlinks verwendet. Dies ist die Grundlage für das WWW. Die Programme, die die Struktur und den logischen Aufbau des Dokuments interpretieren und als formatierte Webseiten (eventuell mit Interaktionselementen wie Links oder Formularen) darstellen, werden Webbrowser genannt. Dem Text wird durch Auszeichnung (engl. *mark-up*) von Textteilen mit in der Regel paarweisen (öffnenden und schließenden) Tags eine Struktur und Logik verliehen. Die jeweils zusammengehörenden Tags bilden zusammen mit dem Inhalt zwischen den Tags ein Element. Diese Elemente lassen sich nach Regeln, die in einer Dokumenttypdefinition (*DTD*) angegeben sind, verschachteln: `<p>Ein Textabsatz, der ein <em>betontes</em> Wort enthält.</p>`

Neben Elementen mit Start- und End-Tag gibt es auch leere Elemente, wie etwa Zeilenumbrüche oder Bilder: `Eine Textzeile,<br> die hier fortgesetzt wird. <img src="E-Mail-Button.jpg" alt="E-Mail" title="Schreiben Sie mir eine E-Mail!">` (Anmerkungen: zur Übersicht sind die Attribute umbrochen worden, dies ist nicht nötig, aber erlaubt. Ein Zeilenumbruch entspricht im HTML-Quellkode einem Leerzeichen.) Dabei sind diese Tags keine Präsentations-Befehle (obwohl das in Zeiten von HTML 3.2 so war), die dem interpretierenden Webbrowser mitteilen, er müsse beispielsweise bis zum Auftreten des Endtags alles in Fettschrift setzen, sondern eine strukturgebende oder logische Auszeichnung, deren Repräsentation von der Umgebung

abhängig ist. Obwohl HTML-Dokumente in der Regel auf Computerbildschirmen dargestellt werden, kann man sie auch auf anderen Medien ausgeben lassen, etwa auf Papier oder als Sprachausgabe.

Auf die Präsentation in verschiedenen Medien sollte mit HTML kein Einfluss genommen werden, dazu eignet sich CSS vorzüglich. Daher sollten Elemente und Attribute zur Präsentation, wie `<i>...</i>`, `<font>...</font>`, `width`, `...` vermieden werden und stattdessen nur solche verwendet werden, die tatsächlich der Textauszeichnung dienen, wie etwa `...</p>` oder `<em>...</em>`. Die meisten der Präsentations-Elemente und -Attribute wurden in der HTML4-Spezifikation als missbilligt (engl. *deprecated*) markiert. `<H1>Dies ist eine Überschrift</H1>`

In XHTML wurden einige Elemente geändert, damit sie der XML-Syntax entsprechen. Dies betrifft alle leeren Elemente, also solche, die keinen Text umschließen, wie Beispielsweise *br* oder *hr*. Diese werden nach den Regeln der Tag-Minimierung geschrieben als `<br/>` beziehungsweise `<hr/>`. Da einige ältere Browser jedoch Probleme mit dieser Darstellung haben, sollte zur besseren Abwärtskompatibilität ein zusätzliches Leerzeichen eingefügt werden, und das Element wie folgt geschrieben werden: `<br />`. Außerdem sind alle Element-Bezeichner sowie Attribut-Bezeichner klein zu schreiben. Leeren Attributen (z. B. `nowrap`) muss ein Wert zugewiesen werden (z. B. `nowrap="nowrap"`). Text muss immer durch Elemente beschrieben werden und darf nicht *lose* zwischen den body-Start- und -End-Tags stehen.

HTML-Struktur

Ein HTML-Dokument besteht aus drei Bereichen:

1. der Doctype-Deklaration ganz am Anfang der Datei, die die verwendete DTD angibt, z. B. „HTML 4.01 Strict“.
2. dem HTML-Kopf, der hauptsächlich technische oder dokumentarische Informationen enthält, die nicht direkt im Browser sichtbar sind und
3. dem HTML-Körper, der anzuzeigende Informationen enthält.

Somit sieht die Grundstruktur einer Website wie folgt aus:

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN"
  "http://www.w3.org/TR/html4/strict.dtd">
<html>
<head>
<title>Titel der Webseite</title>
Evtl. weitere Kopfinformationen
</head>
<body>
Inhalt der Webseite
</body>
</html>
```

HTML-Kopf

Im Kopf können 6 verschiedene Elemente angewandt werden:

1. **title** bezeichnet den Titel der Seite, der im Browserfenster, sowie oft in Suchmaschinen angezeigt wird,
2. **meta** enthält so genannte Metaangaben (Metatags) und ist weitreichend parametrisierbar,
3. **base** gibt entweder eine Basis-URI an oder einen Basisframe und
4. **link** dient zur Angabe von logischen Beziehungen zu anderen Ressourcen. Am häufigsten zur Einbindung von Stylesheets benutzt.
5. **script** bindet Code in einer bestimmten Skriptsprache ein, hauptsächlich Javascript.
6. **style** beinhaltet CSS-Regeln direkt im Dokument.

HTML-Körper

Eine Hauptüberschrift wird beispielsweise so ausgezeichnet:

```
<h1>Hauptüberschrift</h1>
```

h1 steht für *Heading 1*.

Diese Auszeichnung wird nun als Hauptüberschrift interpretiert. Weiter möglich sind `h2` bis `h6`, Überschriften zweiter bis sechster Ordnung, mit denen sich die Gliederung einer Seite verdeutlichen lässt. Die Präsentation dieser Überschriften ist von ihrer strukturierenden Bedeutung unabhängig und kann mit CSS beeinflusst werden. Auf keinen Fall sollte man die Überschrift-Elemente zur Vergrößerung von Text missbrauchen. Suchmaschinen-Roboter messen Überschriften eine höhere Relevanz als normalem Fließtext.

Beispiel für einen Weblink:

```
<a href="http://www.example.com/">Gehe zu example.com</a>
```

Hier wird auf die Ressource http://www.example.com/ verwiesen. Der Text *Gehe zu example.com* wird dabei als Link dargestellt.

Die logische Beschreibung der Struktur des Textes vereinfacht es zum Beispiel, dass der Text auch einem Sehbehinderten vorgelesen oder als Brailleschrift ausgegeben werden kann. Auch für Suchmaschinen ist das von Vorteil, um HTML-Dateien möglichst sinnentnehmend auszuwerten.

Zur Logik stehen zum Beispiel die Elemente `strong` und `em` bereit, mit denen sich stark hervorgehobener oder betonter Text auszeichnen lässt. Zur visuellen Kenntlichmachung empfiehlt sich die Verwendung der beiden CSS-Regeln

```
strong { font-weight: bold }
em { font-style: italic }
```

die für den Inhalt von `strong`-Elementen und `em`-Elementen die Präsentation in **Fettschrift** bzw. *Kursiv-Schrift* empfehlen.

DTD-Varianten

Beim Entwurf der letzten HTML-Version 4 sollte der Tatsache, dass in viele Webdokumenten noch präsentationsbehaftete Elemente und Attribute eingesetzt werden, Rechnung getragen, aber auch eine stilistisch saubere DTD angeboten werden. Das Ergebnis waren schließlich drei Varianten:

Strict

Diese DTD-Variante war ursprünglich vorgesehen, wurde aber zugunsten der Abwärtskompatibilität fallen gelassen. Es fehlen die meisten Elemente und Attribute zur Beeinflussung der Präsentation (Auszeichnungselemente) und die Syntax muss wohlgeformt sein. Das bedeutet, dass jedes öffnende Element (Tag) ein schließendes Element besitzen muss, die Elementreihenfolge muss beachtet werden und Text muss sich grundsätzlich in einem Containern befinden. Beispielsweise ist Text, der sich frei zwischen den body-Elementen befindet, nicht mehr erlaubt. Diese Restriktionen waren für die Angleichung an XML notwendig und bilden die Grundlage für XHTML.

Transitional

Die Variante, die noch ältere Elemente und Attribute enthält, die auch physisches Markup ermöglichen. Durch diese DTD soll auch Webautoren, die noch nicht auf Trennung von Textauszeichnung und Präsentation umgestellt haben, die Möglichkeit gegeben werden, standardkonformes HTML zu schreiben und soll sicherstellen, dass bestehende nicht standardkonforme Webseiten weitherhin durch aktuelle Webbrowser angezeigt werden können.

Frameset

Diese Variante enthält zusätzlich zu allen Elementen der Transitional-Variante noch die Elemente für die Erzeugung von Framesets. Bei der Verwendung sollte man sich aber an der Strict-Variante orientieren.

Falsche Interpretation von Webdokumenten

Selbige war und ist ein Ärgernis für Webautoren. Auch wenn man standardkonformes HTML schreibt, ist noch lange nicht gewährleistet, dass die entstandenen Webdokumente in den gängigen grafischen Webbrowsern richtig dargestellt werden. Hierbei ist zwar zu beachten, dass eine Anpassung von HTML-Dokumenten an die Gegebenheiten auf der Leserseite durchaus von Vorteil und auch gewollt ist, aber häufig kommt es auch zu wirklichen Falsch-Interpretationen. Durch die ausschließliche und wohlüberlegte Verwendung von strukturierendem und logischem HTML kann in gewisser Hinsicht gewährleistet werden, dass ein Webdokument unter allen Umständen wenigstens zugänglich und nutzbar ist.

Wer aber auf eine bestimmte Präsentation seiner Inhalte Wert legt, wird finden, dass es oft noch an der CSS-Interpretation scheitert. Gerade der am weitesten verbreitete Webbrowser,

der (Internet Explorer) ist zwar sehr gutmütig und lässt so manchen Fehler durchgehen, hat aber in CSS-Belangen einige Defizite, welchen Rechnung zu tragen ist. Die Neuentwicklungen aus der Gecko-Engine (Mozilla, Mozilla Firefox, Netscape, ...), der KHTML-Engine (Konqueror, Safari) sowie aus dem Hause Opera haben hier am wenigsten Probleme, sind aber auch weniger fehlertolerant.

Wenn einem Webautor also wichtig ist, dass Webdokumente zumindest in den gängigen grafischen Webbrowsern konsistent dargestellt werden, kommt er nicht umhin, seine Erzeugnisse in diesen ausgiebig zu testen.

Sprachtyp

Im Gegensatz zu Java oder C ist HTML keine Programmiersprache, auch wenn die Mehrheit der Webbenutzer genau dies annimmt. HTML ist wegen fehlender Elemente wie Wenn-Dann-Bedingungen, Variablen etc. keine Programmiersprache und HTML ist auch keine Seitenbeschreibungssprache wie etwa Postscript weil eben keine Seiten beschrieben werden, sondern Text strukturell und logisch ausgezeichnet wird. Daher ist HTML eine **Textauszeichnungssprache** und wird als solche auch nicht programmiert, gescriptet, gecodet oder gar geproggt. HTML wird schlicht geschrieben.

Ein ähnliches Konzept (logische Beschreibung) wie hinter HTML steht auch hinter dem Satzsystem TeX/LaTeX, das im Unterschied zu HTML jedoch auf die Ausgabe per Drucker auf Papier zielt.

Weiterentwicklungen

Im Laufe der Jahre ist HTML um Elemente erweitert worden, die sich (fast nur) mit der Gestaltung des Dokuments befassen, was der ursprünglichen Idee der Systemunabhänigkeit entgegen lief.

Dazu gehören die Möglichkeiten Seiten dynamisch zu verändern, z. B. mittels JavaScript. Man spricht dann von DHTML (Dynamic HTML). Da diese Möglichkeiten von den Browser-Herstellern (allen voran Microsoft und Netscape) entwickelt wurden, gibt es erhebliche Probleme mit der Interpretation der Elemente zwischen den verschiedenen Browsern.

Eine Rückbesinnung auf die Trennung von Inhalt (Struktur) und Layout wurde durch die Definition der Cascading Style Sheets (CSS) vorgenommen. So soll das Aussehen des Dokuments in einer separaten Datei (dem sogenannte Style Sheet) festgelegt werden. Dies ermöglicht es viel besser das Layout an das Ausgabegerät (oder spezielle Bedürfnisse, z. B. von Sehbehinderten) anzupassen. Diese Vorgabe wird allerdings nicht konsequent umgesetzt.

Als zukünftiges universelles Format wird zunehmend XML (eXtensible Markup Langugage), ein selbstbeschreibendes Datenmodell eingesetzt. Dieses arbeitet ebenfalls mit Tags, die jedoch per DTD (Document Type Definition) sehr restriktiv eingesetzt werden können. Zur Umwandlung von in XML gehaltenen Daten in die entsprechenden „Anzeige-

formate“ kann XSL verwendet werden, eine Style Sheet Sprache zur Transformation z. B. nach XHTML.

- 1992: Erste Version von **HTML** (http://www.w3.org/History/19921103-*hypertext/hypertext/WWW/MarkUp/MarkUp.html*)
- 1997: **HTML 3.2** (*http://www.w3.org/TR/REC-html32*)
- 1999: **HTML 4.01** (*http://www.w3.org/TR/html401/*)

HTML LERNEN

SELFHTML (*http://selfhtml.teamone.de/*) ist ein deutschsprachiges und sehr umfangreiches Projekt, das für viele Themen rund um HTML Referenzmaterial bietet.

Die Lektüre einer modernen linearen Einführung (siehe Weblinks) und die Handarbeit direkt in einem Texteditor ist empfehlenswert, um HTML richtig zu verstehen und voll auszunutzen.

Für schnelle Arbeiten mag die Arbeit in einem grafischen (so genannten WYSIWYG-Editor) zunächst genügen, allerdings produzieren diese Editoren stets ein HTML, das die optischen Vorstellungen widerspiegelt. Strukturelle und logische Informationen, die dem Text erst einen echten Mehrwert geben, lassen sich nur einsetzen, wenn man HTML hinreichend gut verstanden hat. Hinzu kommt, dass diese Editoren oft ungültiges HTML produzieren, was die Darstellung des Dokuments von der Ratekunst des Webbrowsers abhängig macht. Professionelle Webgestalter sind von der Verwendung solcher Editoren mehrheitlich abgekommen.

WEBLINKS

- Frühe Entwürfe von HTML (*http://www.w3.org/MarkUp/historical*) (englisch)
- Das Standardisierungsgremiums W3C (*http://www.w3.org/MarkUp/*) (englisch)
- W3C-Validator zur Überprüfung auf syntaktische Richtigkeit eines HTML-Dokuments (*http://validator.w3.org/*) (englisch)
- Newsgroup *news://de.comm.infosystems.www.authoring.misc/*
- Anleitung zum Erstellen von HTML-Seiten: SELFHTML (http://selfhtml.teamone.de/)

XML

XML (Extensible Markup Language) ist ein Begriff aus der Computertechnik und bezeichnet einen Standard zur Definition von Auszeichnungssprachen, der als vereinfachte Teilmenge von SGML konzipiert wurde.

Gleichzeitig steht XML in einer losen Verwandtschaft zu HTML, welches ursprünglich (d. h. bis einschließlich zur Spezifikationsversion 4.01) selbst als Anwendung von SGML definiert wurde. Mit der „Extensible HyperText Markup Language“ (XHTML) wurde der Übergang zu XML als Definitionsbasis vollzogen. Grund dafür war die einfachere Syntax und damit die Entwicklung einfacherer Parser (die Definition von SGML umfasst 500 Seiten, jene von XML bloß 26).

Die Namen der einzelnen Strukturelemente für eine bestimmte Auszeichnungssprache lassen sich frei wählen, diese Auszeichnungssprachen können dabei alle möglichen Daten beschreiben, als prominentestes Beispiel Text, aber auch Grafiken oder abstraktes Wissen. Ein Grundgedanke hinter XML ist es, Daten und ihre Repräsentation zu trennen. Also beispielsweise Wetterdaten einmal als Tabelle oder als Grafik auszugeben, aber für beide Anwendungen die gleiche Datenbasis im XML-Format zu nutzen.

Fachtermini

Eine XML-Datei, die als **wohlgeformt** bezeichnet wird, muss die Regeln für XML korrekt einhalten (was z. B. Verschachtelungen von Elementen betrifft).

Programme, die XML-Daten verarbeiten, nennt man **XML-Parser**.

Soll XML für den Datenaustausch verwendet werden, so sollten die Daten einer Dokumenttypdefinition oder einem XML Schema entsprechen.

Aufbau einer XML-Datei

Beispiel einer XML-Datei

```
<?xml version="1.0"?>
<enzyklopaedie>
  <eintrag>
    <stichwort>Genf</stichwort>
    <eintragstext>Genf ist der Sitz von...</eintragstext>
    <autor>Hans Wurst</autor>
  </eintrag>
</enzyklopaedie>
```

XML-Dateien sind hierarchisch strukturiert. Die Dateien sind plattformunabhängig, da ihre Kodierung angegeben wird. Der Standard ist Unicode, genauer gesagt UTF-8.

XML-Dokumente können folgende Objekte enthalten:

- Elemente
  - Start-Tag, End-Tag, Empty-Tag
  - Attribute: An einem Element anhängende Schlüsselwort-Werte-Paare
- Entitäten (Platzhalter, die bei der Auswertung durch anderen Inhalt ersetzt werden):

    - Zeichenreferenzen
    - Allgemeine Entitäten (unter anderem externe)
- Verarbeitungsanweisungen
- Kommentare
- CDATA-Abschnitte
- DTD

Dabei sind die Verarbeitungsanweisungen und die Angabe einer DTD mittlerweile obsolet. Eine XML Datei muss genau ein Element in der obersten Ebene enthalten. Unterhalb von diesem können weitere Elemente verschachtelt werden.

Einige Web-Browser können XML-Dateien mit Hilfe eines eingebauten XML-Parsers direkt darstellen. Dies geschieht in Verbindung mit einem Stylesheet. Diese Transformation kann die Daten in ein komplett anderes Format umwandeln, das Zielformat muss nicht einmal XML sein.

Vorgänger von XML

Obwohl der Vorgänger SGML bereits weitaus umfangreicher war, kam es nie zu einer breiten Akzeptanz in der Öffentlichkeit. Der Grund dafür liegt in der Komplexibilität SGML's, die die Softwareentwicklung stark erschwert. Der Bedarf nach einem unbeschränkten weltweiten Informationssaustausch und die Popularität von HTML, brachten das deutlich einfachere XML hervor, das in den letzten Jahren einen starken Boom erfuhr.

APIs zur Verarbeitung von XML

Die Kerntechnologien im XML Umfeld kann man grob aufteilen in: APIs zur Verarbeitung von XML und Sprachen um XML Dateien zu beschreiben.

SAX ist eine standardisierte Möglichkeit, wie eine XML Datei geparst werden kann. Hierbei wird ein Datei-Strom in einen Strom von Ereignissen umgewandelt. Programme können sich für einzelne Ereignisse registrieren, um bei Bedarf ihre Arbeit zu verrichten. Ein Vorteil von SAX ist, dass die gesamte XML Datei nie im Speicher sein muss, das ist aber dann ein Nachteil, wenn man viele Informationen, die über die ganze Datei verstreut sind, zur Verarbeitung benötigt

DOM (Document Object Model) ist der zweite standardisierte Weg, um XML Dateien auszuwerten. Er stellt, wie der Name schon sagt, ein standardisiertes Objektmodell zur Verfügung, mit dessen Hilfe der Inhalt der XML Datei ausgewertet oder manipuliert werden kann. Hierbei ist jedoch die ganze Datei im Speicher, jedoch sind Programme die auf DOM basieren im Allgemeinen einfacher zu verstehen.

Metasprachen

Es gibt zwei standardisierte Möglichkeiten um die Struktur von XML Dokumenten zu beschreiben, Dokumenttypdefinitionen oder mit XML Schema.

DTD

Eine **DTD** (Dokumenttypdefinition) ist eine Beschreibung eines XML Dokuments. Sie wurde zusammen mit XML standardisiert. Mit einer DTD kann allerdings nicht sehr strikt beschrieben werden, wie eine XML Datei aussehen darf. Ein weiterer Nachteil ist die Tatsache, dass die DTD in einer eigenen Sprache abgefasst werden muss.

XML Schema

XML Schema ist die moderne Möglichkeit, XML Dokumente zu beschreiben. Ein Schema ist selbst ein XML Dokument, das es erlaubt komplexere Zusammenhänge als mit einer DTD zu beschreiben.

Java-Applet

Ein **Java-Applet** ist ein Computerprogramm, das in der Programmiersprache Java verfasst wurde. Der Begriff Applet bedeutet soviel wie "little application" (siehe auch Java-Applikation).

Java-Applets wurden eingeführt, damit man Programme für Web-Seiten schreiben kann, die im Webbrowser (auf der Client-Seite) arbeiten und direkt mit dem Benutzer interagieren können, ohne Daten über die Leitung zum Server versenden zu müssen. Davor mussten Interaktionen entweder über CGIs auf der Serverseite bearbeitet werden, oder über Java Script, das jedoch keine vollwertige Programmiersprache darstellt und für umfangreiche Anwendungen nicht geeignet ist.
Java-Applets waren Ende der 1990er Jahre ein Hauptgrund für den Erfolg und die schnelle Verbreitung von Java.
Neben Applets existieren auch Servlets, das sind ebenfalls Java-Programme, die allerdings auf dem Server ausgeführt werden.

Java-Applets im Internetbrowser

Üblicherweise werden Java-Applets von HTML-Seiten aufgerufen.
Um sie ausführen zu können, muss der jeweilige Webbrowser über eine entsprechende Java-VM verfügen. Diese VM, die Laufzeitumgebung für das Java-Applet, kann entweder Teil des entsprechenden Browsers sein, oder in Form eines Plugins nachträglich installiert werden (z.B. JRE 1.4 von Sun, siehe Weblinks).

Internetbrowser mit integrierter Java-VM sind zum Beispiel :
(Der Funktionsumfang dieser VMs entspricht nur der JDK-Version 1.1, sie sind aus heutiger Sicht im Grunde veraltet eine aktuelle VM lässt sich aber nachinstallieren, siehe Weblinks)

- Internet Explorer von Microsoft (Aktuell plant Microsoft die Java-Unterstützung komplett aus dem IE zu entfernen. Um die Frage der Integration von Java in den Internet Explorer bzw. in Microsoft Windows prozessiert Sun in den USA gegen Microsoft)
- Netscape Navigator 3.x und 4.x

Internetbrowser ohne integrierte Java-VM sind zum Beispiel :
(Um mit diesen Browsern Applets nutzen zu können muss eine VM nachinstalliert werden, siehe Weblinks)

- Mozilla und alle auf dessen Code basierenden Browser, wie Netscape Navigator 6.x und 7.x oder Mozilla Firefox
- Opera

Applet Programmierung

Die Einbindung des Applets in den HTML-Code von Webseiten erfolgt mit dem OBJECT-Tag, dem APPLET-Tag oder dem EMBED-Tag.
Alle Java-Applets werden von der Java-Klasse Java.applet.Applet abgeleitet.
Sie verfügen unter anderen über die Methoden init() start() stop() und destroy() besitzen jedoch keine main() Methode.

Java sowie die nötigen Erweiterungen für die Webbrowser finden sich unter http://java.sun.-com

Cookie

Ein **Cookie** (am. englisch *Plätzchen, Keks*) bezeichnet Informationen, die ein Webserver zu einem Browser sendet, um dem zustandslosen HTTP-Protokoll die Möglichkeit zu geben, Information zwischen Aufrufen zu speichern.

Funktionsweise

Cookies werden in den Header-Teilen von HTTP-Anfragen und HTTP-Antworten übertragen.

Man kann zwischen *persistenten Cookies* und *Session-Cookies* unterscheiden. Erstere werden dauerhaft gespeichert (z. B. auf der Festplatte), während Letztere nur für die Länge einer Sitzung gespeichert werden.

Wenn ein Webserver Cookies zu einem Webbrowser sendet, werden sie lokal auf dem Endgerät gespeichert (auf Computern üblicherweise in einer Textdatei). Die Cookies werden dann bei jedem Aufruf der entsprechenden Website zum Server übertragen. Damit ist eine beständige Verbindung zwischen dem Browser und Server gewährleistet. Cookies können beliebige Informationen enthalten, die ausschließlich vom Webserver bestimmt werden. Dieses Konzept wurde von Netscape entwickelt und ist in **RFC 2109** spezifiziert.

Verwendung

Cookies werden z. B. von Google dazu verwendet, um persönliche Einstellungen zu speichern. Damit ist es möglich, diese Website zu besuchen, ohne jedesmal die Einstellungen erneut vornehmen zu müssen. Auch Online-Shops verwenden Cookies, um so genannte virtuelle Einkaufskörbe zu ermöglichen. Der Kunde kann damit Artikel in den Einkaufskorb legen und sich weiter auf der Website umschauen, um danach die Artikel zusammen online zu kaufen. Cookies dienen auch der Sicherheit. Da man sich bei manchen Websites wie Wikipedia per Passwort einloggen kann, werden Cookies gesetzt, um genau diesen Nutzer eindeutig zu erkennen und damit nicht bei jedem Aufruf der Website das Passwort erneut eingeben werden muss. Das ist sicherer und weniger aufwändig als z. B. Session-IDs (Zahlenfolge, die nur für eine Session gültig ist) im URI.

Gefahren

Die eindeutige Erkennung kann allerdings von Firmen ausgenutzt werden. Cookies werden dabei dazu verwendet, um Benutzerprofile über das Surfverhalten zu erstellen. Ein Online-Shop kann z. B. diese Daten mit dem Namen des Kunden verknüpfen (wenn man Kunde bei ihm ist) und zielgruppengerechte Spam-Mails schicken. Marketingfirmen, die bei vielen Websites Werbebanner haben, können mit Cookies sogar über einzelne Websites hinweg den Benutzer verfolgen.

Erlauben oder Sperren?

Einen Kompromiss zwischen den Vor- und Nachteilen von Cookies kann erzielt werden, indem man seinen Browser so konfiguriert, dass persistente Cookies nicht zugelassen (erschwert z. B. die Erstellung von Benutzerprofilen) und Session-Cookies zugelassen (z. B. für Webeinkäufe, Passwörter) werden. Außerdem bieten die meisten Browser die Möglichkeit, Cookies selektiv für bestimmte Domänen zu erlauben bzw. zu sperren oder nach dem surfen automatisch zu löschen.

Common Gateway Interface (CGI)

Die **CGI**-Schnittstelle (Common Gateway Interface - Allgemeine Vermittlungsrechner-Schnittstelle) ist eine Möglichkeit, Programme oder Scripts im Web bereitzustellen, die von

HTML-Dateien aus aufgerufen werden können, und die selbst HTML-Code erzeugen und an einen Web-Browser senden können. CGI ist also eine, schon länger bestehende, Variante Webseiten dynamisch bzw. interaktiv zu machen.

Um die CGI-Schnittstelle zu verwenden, muss diese von der Web-Server-Software unterstützt werden. Dabei ist wichtig, dass diese Software dem Programm/Script immer 3 Dinge zur Verfügung stellt:

- Umgebungsvariablen (z. B. SERVER_NAME), deren Inhalte dem Programm helfen sich „vor Ort“ zu orientieren und über aktuelle Einstellungen zu informieren.
- Weiterleitung von Ausgaben, meistens als dynamisch erzeugte HTML-Seite (oder Seitenteile), aber auch als Einträge in Fehlerprotokolldateien.
- Einholen von Formulareingaben oder Aufrufparametern z. B. aus HTML-Seiten, damit das CGI-Programm/-Script auf diese reagieren kann. Dabei können solche Daten als Umgebungsvariable (GET-Methode) oder über einen Eingabe-Kanal (POST-Methode) Eingang ins Programm/Script finden, wobei letztere Möglichkeit sicherer ist.

Wie diese Daten strukturiert sind, ist die eigentliche Schnittstellenbeschreibung (deshalb „interface“).

CGI Programme können also in allen möglichen Programmiersprachen geschrieben sein. Es muss lediglich auf dem Server ein entsprechender Laufzeitinterpreter vorhanden sein, oder das Programm muss für das Serverbetriebssystem kompiliert worden sein.

Am weitesten verbreitet ist hierbei Perl.

Ein Nachteil, der der CGI-Ausführung nachgesagt wird, ist dass sie langsamer sei als andere Möglichkeiten (s. z. B. Servlet) da für jeden CGI-Aufruf eine neue Programm-Instanz ausgeführt werden muss. Deshalb wird CGI heutzutage nicht mehr so oft eingesetzt, denn selbst Ansätze wie FastCGI, welches gewisse Nachteile von CGI aufhebt, konnten sich zumindest nicht auf breiter Front durchsetzen. Andererseits wird dieser Nachteil von modernen Webserver-Typen für einige Programmiersprachen weg optimiert (z. B. bietet der weltweit meistbenutzte Webserver Apache mit dem Modul mod_perl die Möglichkeit, einen Perl-Interpreter in den Webserver selbst einzubetten, was – neben anderen Vorteilen – die Ausführungsgeschwindigkeit massiv erhöhen kann). Alle derartigen Lösungen sind jedoch nicht mehr programmiersprachen-unabhängig.

Weitere serverseitige Technologien sind ASP, ColdFusion, JSP/Servlet, PHP

PHP

PHP (rekursives Akronym für "**PHP**: **H**ypertext **P**reprocessor", ursprünglich "**P**ersonal **H**ome **P**age Tools") ist eine Programmiersprache, die hauptsächlich zur Erstellung dynamischer Webseiten verwendet wird. Es handelt sich um eine Skriptsprache, mit einer an C bzw. Perl angelehnten Syntax.

PHP zeichnet sich besonders durch die leichte Erlernbarkeit, breite Datenbankunterstützung und Internet-Protokolleinbindungen, sowie die Verfügbarkeit zahlreicher, zusätzlicher Funktionsbibliotheken aus. Es existieren zum Beispiel Bibliotheken, um allein mit PHP GTK-Anwendungen zu entwickeln.

PHP ist eine serverseitig interpretierte Sprache. Das bedeutet, dass im Gegensatz zu HTML oder Javascript der Quelltext nicht direkt an den Browser übermittelt, sondern erst vom Server ausgeführt wird. Die Ausgabe des Skriptes wird dann an den Browser geschickt. Die Ausgabe wird in den meisten Fällen eine HTML-Seite sein, es ist aber auch möglich, mit PHP andere Datentypen wie z. B. Bilder zu generieren.

Die Vorteile der serverseitigen Ausführung sind, dass beim Clienten (Browser) keine speziellen Fähigkeiten erforderlich sind oder Inkompatibilitäten auftreten können, wie es z. B. bei Javascript und den verschiedenen Browsern der Fall ist. Außerdem bleibt der PHP-Quelltext der Seite auf dem Server, nur der generierte Code ist für den Besucher einsehbar. Gleiches gilt für andere Ressourcen wie z. B. Datenbanken, die daher auch keine direkte Verbindung zum Clienten benötigen.

Nachteilig ist, dass jede Aktion des Benutzers erst bei einem erneuten Aufruf der Seite erfasst werden kann. Außerdem wird jede PHP-Seite vom Server interpretiert, wodurch die Auslastung des Servers steigt. Diese Vor- und Nachteile sind nicht PHP spezifisch, sondern treten bei grundsätzlich jeder Webapplikation auf.

PHP ist zeitweise etwas ungesteuert gewachsen, so funktioniert der Zugriff auf eine Datenbank mittels der MySQL-Funktionen anders als über ODBC; noch deutlicher wird dies beispielsweise bei Inkonsistenzen der Funktionen zur String-Bearbeitung. Zwar besitzt PHP bereits seit Version 3 grundlegend die Funktionalität, um objektorientiertes Programmieren zu unterstützen (diese wurden in Version 4 deutlich verbessert), bisher ist jedoch die gesamte Standardbibliothek prozedural angelegt. Auch bei objektorientierten Sprachen übliche Features wie Kapselung der Daten (z. B. private Variablen), Destruktoren (ersatzweise lässt sich aber in den meisten Fällen die Funktion register_shutdown_function() verwenden) oder Fehlerbehandlung per Exceptions sucht man in PHP 4 noch vergeblich.

Mit der zur Zeit im Beta-Status befindlichen Version PHP 5 soll dieser Missstand behoben werden, dafür sind aber Änderungen am Verhalten von PHP selbst nötig (so werden Variablen in PHP 5 wie bei den meisten Sprachen nur noch eine Referenz auf das Objekt enthalten und nicht wie in PHP 3 und 4 das Objekt selbst).

Entwicklung der Sprache

PHP wurde 1995 von **Rasmus Lerdorf** entwickelt. PHP stand damals noch für Private Home Page Tools und war ursprünglich eine Sammlung von Perl-Skripten. Bald schrieb er jedoch eine größere Umsetzung in C, worin PHP auch heute noch geschrieben ist. Das schließlich veröffentlichte PHP/FI (FI stand für Form Interpreter) war Perl sehr ähnlich, wenn auch viel eingeschränkter, einfach, und ziemlich inkonsistent.

PHP 3 wurde von **Andi Gutmans** und **Zeev Suraski** neu geschrieben, da das inzwischen erschienene PHP/FI 2 ihrer Meinung nach für eCommerce zu schwach war. Auch die Bedeu-

tung des Akronyms „PHP“ wurde hierbei geändert. Lerdorf kooperierte mit Gutmans und Suraski und die Entwicklung von PHP/FI wurde eingestellt. Die von Gutmans und Suraski gegründete Firma Zend entwickelte in der Folge die Zend Engine 1, die das Herz von PHP 4 darstellt.

Da das World Wide Web Ende der 1990er Jahre stark wuchs, bestand großer Bedarf an Skriptsprachen, mit denen sich dynamische Webseiten realisieren ließen. PHP wurde mit der Zeit populärer als Lösungen wie der vorherige De-facto-Standard Perl, welches mit der extrem flachen Lernkurve von PHP nicht konkurrieren konnte, oder Microsofts ASP.

Lizenz

PHP wird unter der **PHP License** vertrieben, einer Softwarelizenz, die die freie Verwendung und Veränderung der Quelltexte erlaubt. Die Software kann somit kostenlos aus dem Internet geladen werden; daneben ist PHP jedoch schon bei einigen Betriebssystemen (z. B. bei allen relevanten Linux-Distributionen oder Mac OS X) im Lieferumfang enthalten. Neben der aktuellen, stabilen Version 4 kann man auch die Version 5 zum Testen von der offiziellen Homepage herunterladen.

Weblinks

- Offizielle Website (http://www.php.net) mit deutschsprachigem Handbuch
- PHP-Center (http://www.php-center.de) – Deutschsprachiges Portal
- PHP-Homepage (http://www.php-homepage.de) – Deutschsprachiges Portal

Webserver

Ein **Webserver** ist ein Server-Programm, welches Dateien zur Verfügung stellt. Diese Dateien sind normalerweise Webseiten, Bilder und Stylesheets. Für den Webserver macht es aber keinen Unterschied, welche Art von Dateien er ausliefert.

Jedes Mal, wenn eine Webseite angefordert wird (beispielsweise durch Anklicken eines Links), wird eine HTTP-Anfrage an einen Webserver gerichtet. Dieser Webserver wird dann die gewünschte Seite zurücksenden. Standard-Port für den Webserver ist 80.

Üblicherweise werden alle Seitenanfragen in einem Logfile abgespeichert, aus dem man mittels Logfile-Analyse unterschiedliche Statistiken über die Besucher bzw. Zugriffe generieren kann.

Die wichtigste Webserver Software ist der Apache Server (Open Source) mit einem Marktanteil von 67% (2004) und seit 1996 ununterbrochen Markführer. Daneben gibt es noch den **Internet Information Server** von Microsoft mit etwa 21% und SunONE mit 3%. Fast völlig verschwunden ist NCSA von Netscape der bis 1996 Marktführer war aber schon 1998 im Promille-Bereich landete.

LAMP

LAMP ist eine Abkürzung für den kombinierten Einsatz der Softwareprodukte **Linux, Apache, MySQL** und wahlweise **PHP** oder **Perl** sowie neuerdings auch immer öfter Python. (O'Reillys Webseite ONLamp.com (http://onlamp.com/) erklärt deswegen auch gerne, das P stünde für PHPerlthon.) Meistens geht es dabei um die Erstellung dynamischer WebSites.

Das Pendant für Windows heißt dementsprechend WAMP.

Allgemein heißt dies, dass unter einem bestimmten Betriebssystem ein Webserver läuft, der Text-Dateien mit Skriptsprachen-Quelltexten enthält, die bei Aufruf durch einen Webbrowser zumindest teilweise Daten aus einer Datenbank auslesen, und aus diesen zur Laufzeit Webseiten generieren.

Barrierefreies Internet

Barrierefreies Internet bezeichnet Internet-Angebote, insbesondere Webseiten, die von behinderten Menschen ohne Probleme genutzt werden können.

Statistisch gesehen sind Menschen mit Behinderung überdurchschnittlich häufig im Internet. Doch die besonderen Bedürfnisse behinderter Menschen werden hier kaum berücksichtigt. Es ist zu wenig bekannt, dass sich blinde und sehbehinderte Nutzer Webseiten per Software vorlesen oder in Braille-Schrift ausgeben lassen. Bei der Gestaltung von Webangeboten (Webdesign) wird die Farbgebung zumeist gewählt, ohne auf Menschen mit Rot/Grün-Sehschwäche Rücksicht zu nehmen; Schaltflächen und Navigations-Elemente sind für Menschen mit motorischen Schwächen kaum zu erreichen und Sehbehinderte sind bei einer Navigation aus Bildern oder gar Java oder Flash-Elementen bestehend benachteiligt. Sie alle benötigen aber Internet-Angebote, die ihren besonderen Bedürfnissen gerecht werden.

Europäische Union

In der EU gibt es 37 Millionen Menschen mit Behinderungen. Der Anteil älterer Menschen an der Gesamtbevölkerung nimmt stetig zu; derzeit sind ca. 20 Prozent der Bevölkerung über 60 Jahre alt. Die e-Europe-Initiative (Dezember 1999) zur Informationsgesellschaft benennt als eines von zehn Zielen die Teilhabe aller, ungeachtet von Alter und Behinderung. Der Aktionsplan zu e-Europe benennt hierfür unter anderem die Vorhaben: Einführung der Richtlinien der Web Accessibility Initiative (WAI) (WAI englisch) bis 2002 in der öffentlichen Verwaltung und Design-for-All-Standards bis 2003. In Bezug auf das erste Vorhaben hat die Bundesregierung bereits gehandelt:

DEUTSCHLAND

In Deutschland gelten acht Millionen Menschen als behindert. Vier von fünf Menschen mit Behinderungen nutzen inzwischen das World Wide Web. Bereits im Juni 1996 wurde der Artikel 3 des Grundgesetzes geändert:

> *„Niemand darf wegen seiner Behinderung benachteiligt werden."*

Zum 1. Mai 2002 ist das „Gesetz zur Gleichstellung behinderter Menschen und zur Änderung anderer Gesetze" (Behindertengleichstellungsgesetz - BGG) vom 27. April 2002 in Kraft getreten. Ziel des Bundesgesetzes ist es,

> *„die Benachteiligung von behinderten Menschen zu beseitigen und zu verhindern sowie die gleichberechtigte Teilhabe von behinderten Menschen am Leben in der Gesellschaft zu gewährleisten und ihnen eine selbstbestimmte Lebensführung zu ermöglichen." (§ 1, Gesetzesziel).*

In diesem Gesetz hat der Bund Regelungen zur Herstellung von Barrierefreiheit für seine Verwaltung gesetzt, die auch die Informationstechnik betreffen. Dabei verpflichtet sich die Bundesverwaltung unter anderem, ihre Internet- und Intranet-Angebote grundsätzlich barrierefrei zu gestalten.

Eine entsprechende Rechtsverordnung (Barrierefreie Informationstechnik Verordnung - BITV) von Bundesinnenministerium und Bundesministerium für Arbeit und Sozialordnung regelt die Maßgaben hierfür. Die Anlage 1 der Rechtsverordnung enthält keine Vorgaben zur grundlegenden Technik (Server, Router, Protokolle), sondern listet Anforderungen auf, die sich an den Richtlinien der WAI orientieren. Der Bund führt zwei Prioritäts-Stufen mit insgesamt 28 Anforderungen und über 60 zu erfüllende Bedingungen auf. Für die Anpassung bestehender Angebote ist eine Übergangsfrist bis zum 31. Dezember 2005 vorgesehen; neue Angebote haben die Regelungen sofort zu berücksichtigen.

Im Aktionsbündnis für barrierefreie Informationstechnik AbI (*http://abi-projekt.de*) haben sich Behindertenverbände, Forschungseinrichtungen, und andere zusammengeschlossen, um die Umsetzung der Barrierefreieheit im Internet zu fördern. AbI bietet auf dem Informationsportal WOB11 (*http://wob11.de*) Informationen zum Thema barrierefreies Internet.

Die Aktion Mensch und die Stiftung Digitale Chancen haben am 3. Dezember 2003 in Berlin die besten deutschsprachigen, barrierefreien Websites mit dem BIENE-Award ausgezeichnet.

USA

Nach vorsichtigen Schätzungen gelten 39,1 Millionen US-Amerikaner (15 Prozent der Bevölkerung) als behindert. Die USA sind bezüglich der Einführung der Barrierefreiheit in der öffentlichen Verwaltung auf Bundes- und Einzelstaats-Ebene Vorreiter: Bereits 1990 wurde mit dem Americans with Disabilities Act (ADA) ein Behindertengleichstellungsgesetz erlassen, dessen Umsetzung vom Bundes-Justizministerium überwacht wird. Der 1998 erweiterte Abschnitt 508 des *Rehabilitation Act* bindet alle Bundesbehörden bezüglich ihrer

Informations-Angebote. Die hier durch eine unabhängige Bundeseinrichtung erarbeiteten Regelwerke wurden sogar in die Beschaffungsvorgaben aufgenommen und müssen von allen Firmen erfüllt werden, die an die Regierung Waren oder Dienstleistungen verkaufen.

Die meisten Einzelstaaten bieten ihre Internet-Angebote alternativ auch in einer „nur Text"-Version an oder erfüllen, wie beispielsweise Delaware, bereits vollständig alle Priorität-1-Anforderungen der WAI. Die eGovernment-Leitstelle von Delaware ist unter anderem damit befasst, die Umsetzung der WAI-Richtlinien bei allen Verwaltungs-Angeboten des Staates zu befördern. Auch im kommunalen Bereich gibt es Beispiele. So erfüllt der Internet-Auftritt der Stadt Orlando (Florida) ebenfalls die WAI-Vorgaben der Priorität 1.

WEBLOG

Weblog ist ein Kunstwort aus 'Web/WWW' und 'Logbuch'.

In einem **Weblog** (auch *Blog*) hält ein Autor ('Blogger') seine 'Surftour' durch das Internet fest, indem er zu besuchten Seiten einen Hypertext-Eintrag schreibt. Er linkt auf andere Webseiten und kommentiert aktuelle Ereignisse oder äußert Gedanken und Ideen. Viele Einträge bestehen aus Einträgen anderer Weblogs oder beziehen sich auf diese, so dass Weblogs untereinander stark vernetzt sind. Die Gesamtheit aller Weblogs bildet die Blogosphere.

Neue Einträge stehen in einem Weblog chronologisch gereiht an oberster Stelle. Weblogs sind vergleichbar mit Nachrichtenseiten oder Newslettern, jedoch persönlicher - sie selektieren und kommentieren. Weblogs sind demnach keine Alternative zu (Online-) Zeitungen, sondern eine Ergänzung. Im Idealfall reagieren Weblogs schneller auf Trends oder bieten weiterführende Informationen bzw. Links zu bestimmten Themen. Die meisten Weblogs haben eine Kommentarfunktion, die es den Lesern ermöglicht, einen Eintrag zu kommentieren und so mit dem Autor oder anderen Lesern zu diskutieren.

Mit dem stetigen Wachsen der „Weblog-Welt" nimmt auch die Vielfalt an unterschiedlichsten Weblog-Formen zu. So gibt es weiterhin die „klassischen" Weblogs, aber auch eine wachsende Zahl persönlicher Tagebücher, die als Weblog geführt werden und sich vor allem deren einfach zu bedienende Technik zu Nutze machen. Etliche Weblogs enthalten eine Mischung aus Kommentaren, Netzfunden und Tagebuch-Einträgen und dienen in erster Linie der Unterhaltung oder der persönlichen Selbstdarstellung im Internet.

Charakteristisch für Weblog-Software ist, dass sie es auch einem unerfahrenen Nutzer möglich macht, Webseiten zu publizieren. Weblog-Systeme sind einfache Content Management Systeme.

Zur Suche in Weblogs gibt es spezielle Suchmaschinen, z. B. Feedster (*http://www.feedster.com/*) oder blogg.de (*http://www.blogg.de*)

Literatur

- Sven Przepiorka: Diplomarbeit zu Weblogs und deren technische Umsetzung (*http://www.tzwaen.com/publikationen/diplomarbeit.php*), 2003
- Rebecca Blood: *The weblog handbook: practical advice on creating and maintaining your blog*. 2002. ISBN 0-73820756-X
- Aktuelle Liste von Büchern zum Thema: http://www.bloghaus.net/buch.php (*http://www.bloghaus.net/buch.php*)

Wiki

Wikis, auch **WikiWikis und WikiWebs,** sind im World Wide Web verfügbare Seitensammlungen, die von den Benutzern nicht nur gelesen, sondern auch online *geändert* werden können. Sie sind damit offene Content Management Systeme. Der Name bezieht sich auf *wiki*, das hawaiianische Wort für „*schnell*“.

Wie bei Hypertexten üblich, sind die einzelnen Seiten und Artikel eines Wikis durch Querverweise (Links) miteinander verbunden. Die Seiten lassen sich jedoch sofort am Bildschirm ändern. Dazu gibt es in der Regel eine Bearbeitungsfunktion, die ein Eingabefenster öffnet, in dem der Text des Artikels bearbeitet werden kann.

Um den Text zu formatieren, gibt es meist Zeichenkombinationen, die eine Auszeichnungsart an- und wieder ausschalten. Diese sogenannten *Tags* werden im Eingabefenster an entsprechender Stelle eingegeben. In der Wikipedia wird beispielsweise aus der Eingabe „`ein ''kursives'' Wort`“ “ein *kursives* Wort”.

Die Gesamtheit dieser Tags wird als WikiSyntax bezeichnet, und unterscheidet sich je nach verwendeter Wiki-Software. Allen Dialekten ist jedoch zu eigen, dass sie sehr viel einfacher aufgebaut sind als das ansonsten im World Wide Web verbreitete HTML. Diese Beschränkung auf das Wesentliche ermöglicht einer großen Gruppe von Menschen, insbesondere auch Computer-Laien, mit ganz wenig Lern- und Schreibaufwand an diesem System teilzuhaben.

Geschichte

Wikis entstanden als **Wissensmanagement**-Tool im Umfeld der „*Design Pattern*“-Theoretiker. Das erste WikiWeb wurde vom US-amerikanischen Software-Guru Ward Cunningham entwickelt.

Die Welt der Wikis

Von der Wikipedia Tourbus Haltestelle aus kann man eine virtuelle Busrundfahrt durch viele unterschiedliche Wikis unternehmen.

Zur Vernetzung zwischen Wikis gibt es InterWiki, das registrierte Namensräume anbietet (z. B. für Wikipedia oder Wiktionary). Diese haben den Vorteil, dass man aus einem Wiki leichter in ein anderes verlinken kann. Somit bewegt man sich zwischen den Wikis ähnlich komfortabel wie innerhalb eines Wikis.

Ausgewählte Anwendungsbeispiele

- deutschsprachige Wikipedia: http://de.wikipedia.org/
- englischsprachige Wikipedia: http://en.wikipedia.org/
- Das erste WikiWeb: http://c2.com/cgi/wiki?WikiWikiWeb
- DseWiki (http://www.wikiservice.at/dse/wiki.cgi) - Deutsches Software Entwickler Wiki
- deutsche Juristen-Website: (http://www.jurawiki.de)
- GIS-Wiki (http://webgis.dyndns.org:8080/giswiki/Wiki.jsp) - ein Wiki für Geoinformatik im deutschsprachigen Raum

Weblinks

Umfangreiche Listen verschiedener Wikis:

- DasRichtigeWiki (http://www.wikiservice.at/dse/wiki.cgi?DasRichtigeWiki)
- WorldWideWiki.net: OneBigWiki (http://www.worldwidewiki.net/wiki/OneBigWiki) (Englisch)
- BlogsAndWikis (http://199.17.178.148/~morgan/cgi-bin/blogsAndWiki.pl?ExampleWikis) (Englisch)
- http://www.wikiweb.at/ (Deutsch)

Webportal

Der Begriff **Webportal** ist nicht fest zu definieren; allen Definitionsversuchen gemein ist lediglich, dass es sich um eine Website handelt, die versucht, verschiedene regelmäßig benötigte Dienste zu bündeln, oder, eine Übersicht für den Einstieg in einen Themenkomplex zu schaffen. Er wird häuifg auch fälschlicherweise für Webapplikationen benutzt.

Dadurch, dass sehr schnell sehr viele Informationen auflaufen können, die nur schwer sinnvoll darstellbar sind und von spezialisierten Sites besser angeboten werden, leiden viele Portale unter Unübersichtlichkeit und Featuritis. Das und penetrante Werbung führte zu einer negativen Besetzung des Begriffs unter erfahreneren Websurfern.

Webportale als elektronische Märkte

Die Betreiber der Webportale bieten Anbietern und Nachfragern die Möglichkeit, auf effiziente und kostengünstige Art und Weise zu kommunizieren und Ihre Geschäfte schnell und problemlos abzuwickeln. Die Internet-Technologie schafft hierfür die Voraussetzungen.

In der Logistik bieten Webportale oder Marktplätze die Möglichkeiten, Abläufe immer weiter zu optimieren - z. B. ist es so dem Hersteller und dem Handel möglich, auf effiziente Art und Weise zu kommunizieren. Die Basistechnologien des ECR (Efficient Consumer Response) nutzen unter anderem elektronische Marktplätze.

Auch Prozesse wie VMI (Vendor Managed Inventory) nutzen Web-Portale, um die Wiederbevorratung mit einem Minimum an Aufwand zum richtigen Zeitpunkt zu betreiben (bei dem VMI ist der Lieferant verantwortlich für die Bestände seiner Produkte bei dem Händler).

Aus Kundensicht bietet sich eine große Transparenz, um Informationen zu Anbietern und deren Leistungsangeboten zu recherchieren und zu vergleichen. Der Hersteller profitiert von den Möglichkeiten des Internets, seine Zielgruppe genau ansprechen zu können und sich in neue Geschäftsfelder zu etablieren.

Die Webportale bieten beiden Seiten eine Plattform um sich zu begegnen. Die Portale treten also als Vermittler auf und sahen bis vor einigen Monaten ihre Haupteinnahmequelle im Bereich der Werbe- und Vermittlungseinnahmen.

Finanzierung

Wie bereits erwähnt, finanzieren sich die verschiedenen Web-Portale, zum Teil ausschließlich, durch Werbung auf ihrer Seite bzw. durch Werbung in Verbindung mit den verschiedenen Dienstleistungen. Für die Webportale ist eine hohe Besucherzahl von großer Bedeutung, um im immer härter werdenden Wettbewerb überleben zu können. Die Portale wollen, wie jede Unternehmung der Old-Economy, erreichen, dass die Kunden, also die Besucher, immer wieder auf die Portalseite kommen. Um Kunden zu gewinnen, gibt es verschiedene Möglichkeiten. Da es hier unmöglich ist, alle Möglichkeiten der Internetwerbung und die Versuche der Kundenbindung näher darzustellen, werden nur einige Möglichkeiten erwähnt - nicht aber erläutert:

- Softwarebasierte Kundenbindung
    - z. B. die Software von T-Online - Kunden werden mit der Installation automatisch immer wieder auf die Seite geleitet. Viele Nutzer wissen nicht, dass sie die „Startseiten-Einstellungen“ in ihrem Browser ändern können.
    - Zwang oder auch nur Bevorzugung von Software des Internetproviders wie bei AOL oder T-Online.
- Banner: Bild-Animation mit Werbebotschaften eines Unternehmens
- Pop-Up: Werbeeinblendungen in einem neu geöffneten Browser-Fenster

- Brand Flooding: Ein bestimmter Markenname wird immer wieder auf der Web-Seite genannt
- Mailfooter: Werbetexte die unter jeder E-Mail eines Free-Mail-Anbieters zu finden ist – z. B. bei GMX, web.de, T-Online

Andere Dienste und Verbindungsarten

WAP

Das **WAP** (**Wireless Application Protocol**) bezeichnet eine Sammlung von Technologien und Protokollen, deren Zielsetzung es ist, Internetinhalte für die langsame Übertragungsrate und die langen Antwortzeiten im Mobilfunk, sowie für die kleinen Displays der Handys verfügbar zu machen.

Die primäre Aufgabe bei WAP ist es, neben der Berücksichtigung der geringen Displaykapazitäten und Rechenleistung von WAP-Clients, bei der Kodierung der Internetinhalte die offene Struktur und Lesbarkeit einer Auszeichnungssprache (Markup-Language) beizubehalten, und zugleich die Menge der zu übertragenden Daten zu reduzieren. Diese zwei Forderungen stehen zunächst im Widerspruch zueinander:

- eine lesbare Auszeichnungssprache enthält viele Daten, die für die Lesbarkeit notwendig sind (Kommentare, Befehle in lesbarer Form) zur Inhaltsbeschreibung garnicht notwendig sind.
- eine Kodierung in binärer Form erlaubt eine sehr kompakte Umsetzung, weist jedoch weder eine offene Struktur auf, noch ist sie lesbar.

Die Lösung des Problems besteht darin, dass bei WAP zwar die offene Form einer Auszeichnungssprache (bei WAP ist dies die WML - *Wireless Markup Language*) beibehalten wird, diese jedoch nicht als Text, sondern in compilierter Form zum WAP-Client übertragen wird. Dazu erfolgt die Kommunikation zwischen WAP-Client und Web-Server über einen Proxy, den sogenannten WAP-Gateway. Dieser übersetzt die binär vom WAP-Client kommenden Anfragen in Klartext an den Web-Server. Die Antworten des Servers werden im WAP-Gateway compiliert im MIME-Typ WMLC (Wireless Markup Language Ccompiled) an den Client übertragen. Dazu übernimmt der Gateway Aufgaben (wie syntaktische Analyse der WML-Seiten), die im Web der Browser ausführt.

Die Kommunikation zwischen Server und WAP-Gateway verwendet das im Web etablierte Protokoll HTTP. Die Kommunikation zwischen Gateway und WAP-Client erfolgt via WSP (wireless session protocoll). In der Verwendung des Trägers auf der Funkstrecke ist WAP flexibel, möglich sind z. B. CSD (Circuit Switched Data), HSCSD (High Speed CSD), GPRS (General Packet Radio Service), aber auch UMTS (Universal Mobile Telecommunications System). Die Nutzung des GSM-spezifischen USSD-Übermittlungsdienstes (Unstructured Supplementary Service Data)ist ebenfalls möglich.

Der WAP-Standard unterstützt auch einen Push-Service, der es erlaubt, Nachrichten inklusive einer URI ohne Anforderung an den WAP-Client zu senden.

Mit WAP 2.0 wurde das Proxy-Konzept aufgeweicht. Der Standard sieht nun auch die Möglichkeit vor, dass der Client unter Umgehung des Gateways direkt mit dem Web-Server kommuniziert. Damit wird zwar die Abhängigkeit vom einwandfreien Funktionieren das

WAP-Gateways ausgeräumt, allerdings gibt es bereits WAP 2.0-Clients, die WSP nicht mehr unterstützen. Diese können dann die bisher verfügbaren WAP-Gateways nicht mehr verwenden.

I-MODE

i-mode ist ein dem Internet ähnlicher mobiler Dienst für Handys.

Er bietet farbige Texte und Grafiken. Die so genannten i-mode Handsets, das sind i-mode-fähige Mobiltelefone, hatten zu Beginn meist größere Displays als andere Mobiltelefone, spätestens mit der Markteinführung des Siemens S55 in der Wintersaison 2002/2003 ist dies jedoch absolut passee. i-mode wurde von NTT Docomo in Japan entwickelt und hat dort ca. 40 Millionen Nutzer. Damit ist i-mode der erfolgreichste mobile Datendienst der Welt. Außerhalb Japans ist i-mode allerdings nicht so erfolgreich. Im Januar 2004 waren es gerade mal 2 Mio. Nutzer. Kritiker sehen in i-mode eine Totgeburt, weil die Technik in Deutschland zu spät gestartet sei und außerdem UMTS vor der Tür stehe.

Jedoch handelt es sich bei i-mode eher um einen mobilen Datendienst, der unabhängig von der Übertragungstechnik ist. i-mode wird vorraussichtlich sogar durch UMTS weiter verbessert, indem die Inhalte schneller übertragen werden können und somit reichhaltigere Inhalte möglich werden.

i-mode wurde von E-Plus zur CeBit 2002 in Deutschland eingeführt. Den i-mode Dienst gibt es außerem noch in den Niederlanden, Frankreich, Belgien, Spanien, Italien und demnächst in Griechenland.

Zur Darstellung von Internet-Seiten auf i-mode Mobiltelefonen wird iHTML, eine von cHTML (c für compact) abgeleitete Auszeichnungssprache, benutzt. iHTML wurde von der Firma *NTT DoCoMo* entwickelt und ist ein proprietärer Standard. iHTML verwendet eine Untermenge von HTML Tags und ist sehr einfach zu implementieren. Ergänzt werden die tags um einige neue Elemente wie beispielsweise Darstellungssymbole, Tastatursteuerungsbefehle und z. B. Links zu Telefonnummern.

WEBLINKS

- http://www.eplus-imode.de Die offizielle i-mode-Website von E-Plus
- http://www.molipo.de Deutschlands umfangreichstes Linkportal für i-mode Seiten. Die i-mode´-Version ist unter: http://imode.molipo.de verfügbar.

UMTS

Universal Mobile Telecommunications System, besser bekannt unter der Abkürzung **UMTS**, ist ein Mobilfunkstandard der dritten Generation. Ursprünglich vom ETSI stan-

dardisiert, wird UMTS heute vom 3GPP (3rd Generation Partnership Project) weiter gepflegt.

Seit dem 12. Februar 2004 wird UMTS in Deutschland kommerziell vermarktet.

UMTS umfasst erweiterte multimediale Dienste, Satelliten und erdgestützte Sendeanlagen. Folgende Dienste können über UMTS angeboten werden:

- Zwischenmenschliche Kommunikation (Audio- und Videotelefonie)
- Nachrichtendienste (Unified messaging, Video-Sprach-Mail, Chat)
- Informationsverteilung (World Wide Web Browsing, Informationsdienste, öffentliche Dienste)
- Erweiterte Positionsanwendungen (persönliche Navigation, Fahrerunterstützung)
- Geschäftsdienste (Prozessmanagement, Mobilität in geschlossenen Räumen)
- Massendienste (Bankdienste, e-Commerce, Überwachung, Beratungsdienste)

Es wird mehrere Phasen von UMTS geben, die erste Phase (Release 1999) unterscheidet sich vom Vorgängersystem GSM vor allem durch eine neue Funkzugriffstechnologie Wideband-CDMA, die auf CDMA basiert. Durch diese werden höhere Übertragungsraten möglich. Außerdem kann eine Mobilstation (vulgo: das UMTS-fähige Endgerät) mehrere Datenströme gleichzeitig senden bzw. empfangen. Damit können Nutzer beispielsweise gleichzeitig telefonieren und E-Mails empfangen.

Im FDD-Modus (Frequency Division Duplex, Frequenzmultimplex) funken Mobil- und Basisstation in zwei verschiedenen Frequenzbereichen: Im Uplink-Kanal sendet das Mobilteil, im Downlink-Kanal die Basisstation. Derzeit bauen die deutschen UMTS-Netzbetreiber ihre Netze im FDD-Modus auf, die damit erzielbaren Datenraten liegen derzeit bei 384 kbit/s im Downlink.

Im TDD-Modus (Time Division Duplex, Zeitmultimplex) senden Mobil- und Basisstation im gleichen Frequenzband, jedoch zu unterschiedlichen Zeiten. Das Verfahren ist technisch aufwändiger, vor allem wenn sich der Sender bewegt oder weit von der Basisstation entfernt ist, können Timing-Probleme auftreten. Mit W-CDMA im TDD-Modus sollen Datenraten bis zu 2 Mbit/s im Downlink erreicht werden.

Technische Daten

- Grundlegende Funk-Technologie: Wideband CDMA (WCDMA)
- Nutzfrequenzen:
  1. 1900-1920MHz (TDD)
  2. 1920-1980MHz (FDD-Uplink)
  3. 2020-2025MHz (TDD)
  4. 2110-2170MHz (FDD-Downlink)
- Kanalbandbreite: 5MHz

- Nutzertrennung: Code (CDMA); Frequenz (FDMA); Ort/Funkzelle

INTERNET2

Das **Internet2** ist ein Projekt für ein schnelleres Internet basierend auf einem Glasfaser-Backbone. Es wurde 1997 initialisiert von „*The University Corporation for Advanced Internet Development*“ (UCAID).

Das Netz selber erhielt den Codenamen „*Abilene*“ (in Anlehnung an die Stadt Abilene (Kansas), die als Endstation der Eisenbahn im 19. Jahrhundert zum Begriff wurde). Ziel waren ursprünglich Übertragungsraten bis 2.48 Gbit/sec und schneller. Seit 2004 stehen 10 Gbit/sec zur Verfügung. Beachtenswert ist, dass das Internet2 im Gegensatz zum Internet nicht aus dem militärischen Bereich entstanden ist. Die US-Streitkräfte betreiben jedoch verschiedene neuere, weltweite Netzwerke für Forschung, Verwaltung und Kampfeinsatz.

Das Internet2 verbindet seit dem 24. April 1998 115 amerikanische Universitäten und Forschungseinrichtungen. Anfang 2004 waren es bereits mehr als 200 US-Institutionen.

1999 haben auch deutsche, französische, italienische und britische Partnerorganisationen ein Memorandum of Understanding mit der UCAID unterzeichnet, damit sie Anschluss (im wahrsten Sinne des Wortes) erhalten. Dies ist für Deutschland im Jahre 2000 gelungen. Am 30.06.2000 wurde das **deutsche Internet2** eingeführt. Das sogenannte G-Win-Netz ist gröstenteils auf 10-GBit Leitungen Verbindungen aufgebaut und verbindet über 500 deutsche Universitäten und Forschungseinrichtungen.

Internet2 ist eine physische Kommunikationsstruktur und bezeichnet kein explizites Internet_Protocol, es können sowohl IPv4 als auch IPv6 darauf genutzt werden

WEBLINKS

- http://www.ucaid.edu/
- http://www.internet2.org/
- http://www.dfn.de/

SICHERHEIT IM INTERNET

COMPUTERSICHERHEIT

Unter **Computersicherheit** versteht man Maßnahmen, die sicherstellen, dass nur befugte Personen Zugriff auf ein Computersystem haben. Da auf Computersystemen zunehmend wertvollere Daten gespeichert werden, wächst die Wichtigkeit der Computersicherheit. Und da Computersysteme immer komplexer werden, ist absolute Sicherheit in der realen Welt nicht zu erreichen.

Ein Dieb kann erfolgreich die bestbewachte Bank ausrauben. Ein Computerkrimineller kann im bestgesicherten Computer Daten lesen, kopieren und zerstören. Wie in der richtigen Welt ist das Beste, dass man tun kann, es dem Verbrecher möglichst schwer zu machen, die Kosten/Nutzen-Gleichung zu verändern. Die Auswirkungen von Datenverlust kann man reduzieren, indem man Backups anfertigt und sich versichert. Und Sie können das Kosten/Nutzen-Verhältnis verändern, indem Sie nach der Attacke den Verbrecher juristisch verfolgen.

Der einzige Unterschied zwischen Computersicherheit und der realen Welt einer „Banksicherheit“ ist, dass Computersystem oft nur schlecht verstanden werden. Die meisten Menschen haben ein klares Bild von Sicherheitsmaßnahmen wie Zäunen, Mauern, Sicherheitspersonal, Alarmanlagen, Polizei usw. Computersysteme hingegen sind oftmals nicht gegen Datenklau und Zerstörung versichert, damit ist auch „Sicherheitsberatung“ (durch Auflagen von Versicherungen) verloren. Dieser Mangel an Versicherung ist angesichts der Wichtigkeit eines möglichen Verlustes schon bemerkenswert. Dies stammt vom selben Mangel an Wissen, obwohl die Gründe komplexer sein mögen.

Ein Jugendlicher, der im Kaufhaus umher spaziert und sich eine Trophäe nimmt, um sie seinen Freunden zu zeigen, wird in der realen Welt nicht wie ein gefährlicher Krimineller behandelt. Sollte so ein Entdecker in das Computersystem einer Firma einbrechen, sollte das Management schwere Geschütze auffahren, der Eindringling riskiert eine Bestrafung, falls er gefasst wird. Der Wissensmangel ist das größte Risiko in einer Firma. Das für die Technik zuständige Personal wird sich eher um die Hard- und Software als um die Ausbildung der Benutzer kümmern. Vor allem die Gefahren des Social Engineering werden oft ignoriert.

Natürlich sind die Parallelen zwischen Computersicherheit und Sicherheit in der Realen Welt aus einer Reihe von Gründen nicht exakt. z. B. ist Vandalismus sehr viel gefährlicher in einem Computersystem, weil es potentiell viel zerstörerischer ist. Ein Vandale kann in tausenden Computersystemen rund um die Welt Verwüstung anrichten, ohne allzu große Gefahr, erwischt zu werden. Natürlich wäre das Ergebnis nicht so sichtbar wie ein Graffiti, aber alles in allem ist es doch erstaunlich, dass die Wahrscheinlichkeit für einen Schaden durch einen Computervirus oder einen Computerwurm heute klein ist.

Heutzutage besteht Computersicherheit in der Hauptsache aus präventiven Maßnahmen wie beispielsweise Firewalls. Eine Firewall können wir uns vorstellen als einen Zaun um unser Lagerhaus. Dies ist ein erster Schritt. Aber nicht genug, falls man den Zaun nicht bewacht oder falls man jedem einen Schlüssel gibt, der darum telefonisch bittet (Social Engineering). (...) Jedoch werden viele Computersysteme nicht überwacht, und die Zahl der Computerkriminellen, die zur Rechenschaft gezogen werden, ist sehr gering. Somit ist es kein Wunder, dass in dieser Situation kaum jemand versichert ist: die Polizei wäre sehr teuer.

Um es kurz zu machen, der Mangel an Computersicherheit ist eine vielschichtige Bedrohung, die nur durch eine vielschichtige Abwehr beantwortet werden kann. Der Kauf einer Software aus dem Regal ist kein Ersatz für eine umsichtige Untersuchung der Risiken, der möglichen Verluste, der Abwehr und der Sicherheitsbestimmungen auf einer hohen Ebene der Firma.

Literatur

- Claudia Eckert: *IT Sicherheit*, Oldenbourg , ISBN 3-486-27205-5

Firewall

Als **Zugangsschutzsystem** oder **Firewall** bezeichnet man ein organisatorisches und technisches Konzept zur Trennung von Netzbereichen, dessen korrekte Umsetzung und dauerhafte Pflege. Ein oft benutztes Instrument der Umsetzung ist ein Stück Hardware, das zwei physisch getrennte Netzbereiche genau so verbindet, wie es im Konzept zugelassen wird. Dieses Stück Hardware bezeichnet man als Firewall-Rechner/System oder verkürzt als Firewall.

Umgangssprachlich ist mit einem Zugangsschutzsystem sehr oft die Software gemeint, welche den Datenverkehr zwischen den getrennten Netzbereichen kontrolliert und regelt. Man muss also zwischen dem Konzept Zugangsschutzsystem, und den zwei Hauptbestandteilen des Zugangsschutzsystems, nämlich Hardware und Software, unterscheiden.

Ein Zugangsschutzsystem arbeitet auf den Schichten 2 bis 7 des OSI-Referenzmodells. Benutzt man solch ein Zugangsschutzsystem innerhalb einer MAC-Layer-Bridge, so könnte sie Multicast- und Broadcast-Pakete aus dem Netzwerkverkehr heraus filtern und dadurch die Ausbreitung von Broadcaststürmen verhindern.

Der häufige Einsatz eines Zugangsschutzsystem besteht darin, den Verkehr zwischen einem lokalen Netzwerk und dem Internet zu kontrollieren und zu steuern. Ein komplexes Szenario stellt die DMZ dar.

Ein softwarebasierter Teil eines Zugangsschutzsystems in diesem Sinne ist der Paketfilter. Er hat die Aufgabe bestimmte Filterungen oder Reglementierungen im Netzwerkverkehr vorzunehmen. Dieser Paketfilter definiert Regeln, welche festlegen, ob einzelne oder zusammenhängende Pakete das Zugangsschutzsystem passieren dürfen oder abgeblockt werden.

Wenn man sich das Internet als eine gigantische Ansammlung von Häusern vorstellt, dann stellen die IPs sozusagen die Hausnummern dar. (Straßennamen sind in der Welt des Internets unbekannt.) Unter einer bestimmten Hausnummer kann man nun direkt mit einem Rechner kommunizieren, egal wo sich dieser Rechner befindet. In den einzelnen Etagen dieser Rechner wohnen nun die verschiedenen Dienste wie HTTP, FTP oder SSH. Ein Zugangsschutzsystem kann nun verschiedene Etagen für die Besucher aus dem Internet sperren, d. h. jede Verbindung aus dem Internet wird an der Haustüre schon abgewiesen. Durch die entsprechende Konfiguration eines Zugangsschutzsystems kann so ein Computernetzwerk vor Angriffen und/oder Zugriffen geschützt werden.

Personal Firewalls oder auch Desktop Firewalls sind Programme, die lokal auf dem zu schützenden Rechner installiert sind. Ihre Wirkung ist allerdings fraglich: Ist der Rechner ordentlich konfiguriert und laufen vertrauenswürdige Programme, so wird das System selbst nur sinnvolle Pakete annehmen und verschicken. Läuft dagegen zweifelhafte Software auf dem Rechner, die unautorisiert auf das Netz zugreifen will, so wird diese auch soweit gehen, den normalen Weg des Versands zu verlassen und die Personal Firewall umgehen oder ausschalten.

Literatur

- Zwicky, Cooper, Chapman, *Einrichten von Internet Firewalls*, O'Reilly 2001, ISBN 3897211696
- W. R. Cheswick, S. M. Bellovin, A. D. Rubin, *Firewalls and Internet Security - Repelling the Wily Hacker*, 2nd Edition, Addison-Wesley 2003, ISBN 0-201-63466-X

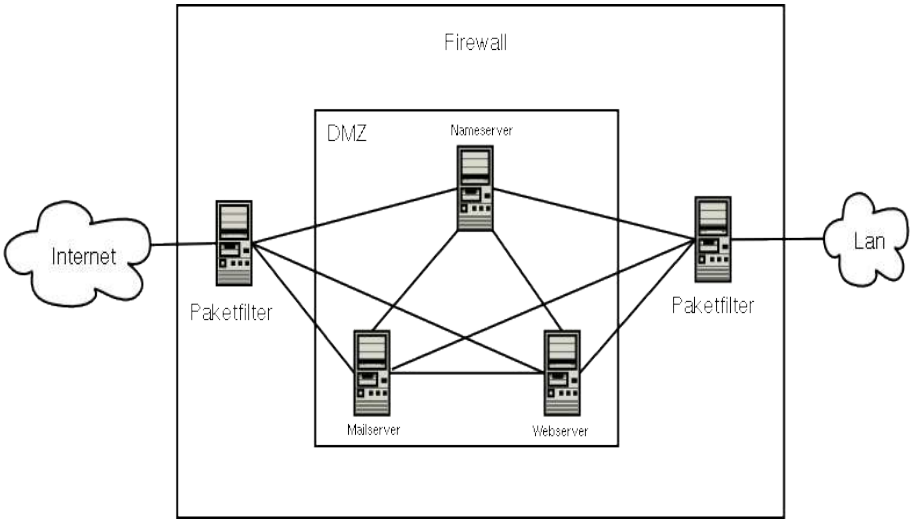
DMZ

Bei einer **DMZ** („Demilitarized Zone“/“entmilitarisierte Zone“) handelt es sich um einen geschützten Rechnerverbund, der sich zwischen 2 Netzwerken befindet. Der Rechnerverbund wird jeweils durch einen Paketfilter gegen das dahinterstehende Netz abgeschirmt.

Der Sinn des ganzen Aufwandes ist es, möglichst auf sicherer Basis Dienste des Rechnerverbundes sowohl dem einem als auch dem anderem Netz zur Verfügung zu stellen. Ein typisches Anwendungsbeispiel ist eine Firma, die einen eigenen Mailserver betreibt.

Dieser Mailserver ist ein Teil der **DMZ** und muss natürlich von außen erreichbar sein, da ansonsten E-Mails nicht zugestellt werden könnten. Anderseits müssen natürlich die Clients, die am LAN angeschlossen sind, ihre E-Mails abholen. Deswegen brauchen auch sie Zugriff auf den Server.

Vorteil einer solchen Lösung ist, dass im Falle einer Kompromittierung eines Servers in der DMZ das interne Netzwerk trotzdem noch geschützt bleibt. Wären die Server nicht in einer DMZ sondern direkt im internen Netzwerk, so wäre auch das gesamte Netzwerk durch eine Kompromittierung betroffen. Gerade weil öffentlich angebotene Dienste oft ein nicht un



erhebliches Angriffsziel darstellen, kann man durch eine **DMZ** das Gesamtrisko erheblich minimieren.

IPSEC

Das **IPsec**-Protokoll wurde 1998 entwickelt, um die Schwächen des Internetprotokolls (IP) zu beheben. Es stellt eine Sicherheitsarchitektur für die Kommunikation über IP-Netzwerke zur Verfügung. Das Protokoll soll Vertraulichkeit, Authentizität und Integrität gewährleisten. Daneben sollen sie vor sogenannte Replay-Angriffen schützen.

IPsec entstand im Zuge der Entwicklung von IPv6 und ist in verschiedenen RfCs spezifiziert:

- RfC 2401: Sicherheitsarchitektur für das Internetprotokoll
- RfC 2402: Authentication Header
- RfC 2406: Encapsulating Security Payload
- RfC 2407: IPsec Domain of Interpration
- RfC 2408: Internet Security Association and Key Management Protocol
- RfC 2409: Internet Key Exchange

Der RfC 2401 bildet das Hauptdokument zu IPsec. Von dort aus werden die oben genannten RfCs referenziert. Wesentliche Inhalte von IPsec sind das Authentication Header (AH)- und das Encapsulated Security Payload (ESP)-Protokoll sowie das Protokoll zum Austausch der Schlüssel.

Im Gegensatz zu anderen Verschlüsselungsprotokollen, wie etwa SSH, arbeitet IPsec auf Netzwerkschicht 3 des OSI-Referenzmodells.

Schlüsselverwaltung (IKE)

Vor dem eigentlichen Start einer verschlüsselten Verbindung muss man sich über die zu verwendenden Schlüssel und Algorithmen klar werden. Hierfür ist IKE gedacht. IPsec arbeitet mit verschiedenen symmetrischen wie asymmetrischen Schlüsseln. Für die Aushandlung der Schlüssel gibt es die Möglichkeit, es manuell oder automatisch zu machen. Für die automatische Schlüsselverwaltung ist das *Internet Key Exchange* (IKE) Protokoll entworfen worden und stellt gleichzeitig den komplexesten Bereich von IPsec dar. IKE basiert auf dem Internet Security Association and Key Management Protokoll, der IPsec Domain of Interpretation, OAKLEY (RfC 2412) und SKEME.

IKE basiert auf UDP und nutzt standardmäßig den Port 500. Es arbeitet in zwei Phasen:

1. Aushandlung einer Security Association (SA) für ISAKMP mittels Aggresive Modus oder Main Modus
2. Erzeugung einer SA für IPsec mittels Quick Modus

Eine Security Association ist ein Vertrag zwischen den kommunizierenden Stellen. Hierin wird festgelegt, welche Authentifizierungs- und Verschlüsselungsalgorithmen genutzt werden sollen.

Main Modus

Der Main Modus kann in der ersten Phase der Internet Key Exchange genutzt werden. Hierbei handeln der Initiator (derjenige, der die Verbindung aufnehmen will) und der Antwortende miteinander SAs für ISAKMP aus. Diese „Verhandlung“ geschieht in folgenden sechs Schritten:

1. Initiator sendet einen oder mehrere Vorschläge mit Authentifizierungs- und Verschlüsselungsalgorithmen
2. Antwortender wählt einen Vorschlag aus und bestätigt
3. Initiator sendet öffentlichen Teil der Diffie-Hellmann-Schlüsselvereinbarung und einen zufälligen Wert (Nonce)
4. Antwortender schickt ebenfalls öffentlichen Teil der Diffie-Hellmann-Schlüsselvereinbarung und einen zufälligen Wert (Nonce)
5. Initiator berechnet Signatur und sendet diese mit seiner Identität an Antwortenden. Diese Daten werden mit einem symmetrischen Schlüssel verschlüsselt.
6. Antwortender schickt gleiche Daten von seiner Seite an den Initiator

Aggressive Modus

Im Aggressive Modus werden die obigen Schritte auf drei zusammengefasst. Hierbei fällt dann die Verschlüsselung des obigen fünften Schrittes weg. Stattdessen werden die Werte im Klartext übertragen. Daher sollte man diesen Modus nach Möglichkeit nicht verwenden.

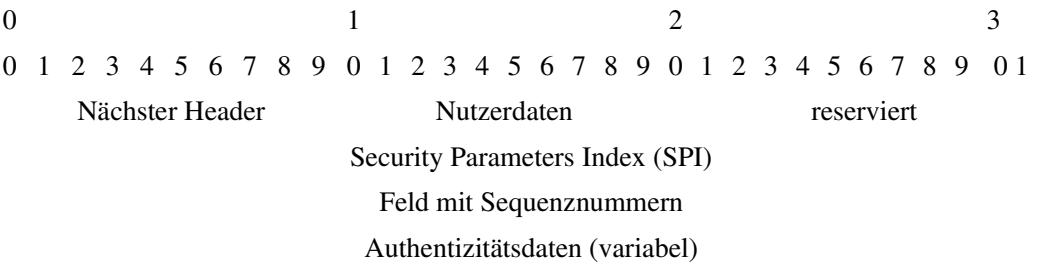
Quick Modus

Der Quick Modus wird in der zweiten Phase von IKE zur Anwendung gebracht. Die gesamte Kommunikation in dieser Phase erfolgt verschlüsselt. Wie in der ersten Phase wird zunächst ein Vorschlag (Proposal) gemacht. Dieser wird zusammen mit einem Hashwert und der Nonce übertragen. Später werden die Schlüssel neu berechnet und es gehen keinerlei Informationen aus den zuvor generierten SAs ein. Dies stellt sicher, dass niemand von der zuvor generierten Schlüsseln auf die neuen schliessen kann.

Authentication Header

Authentication Header (AH) soll die Integrität und Authentizität der übertragenen Pakete sicherstellen. Weiterhin existiert hier ein Schutz gegen Replay-Angriffe. AH versucht alle möglichen Felder eines IP-Datagramms zu schützen. Es werden lediglich Felder ausgeschlossen, die sich während des Laufes eines IP-Paketes ändern können.

Ein AH-Paket sieht folgendermassen aus:



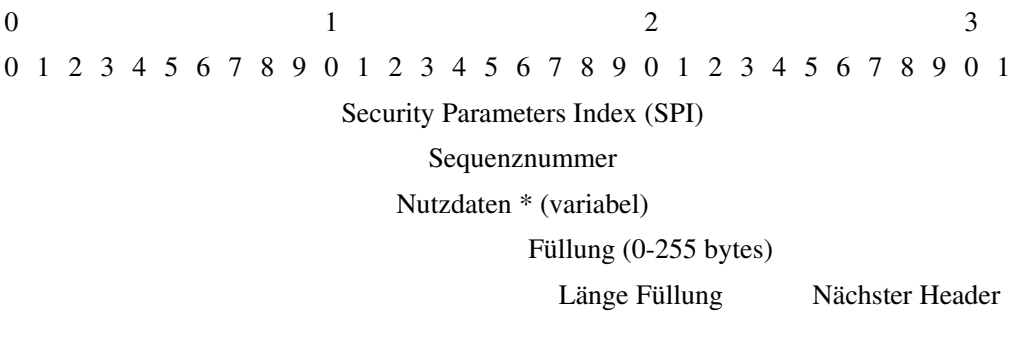
Bedeutung der Felder

- Nächster Header (next header): identifiziert das Protokoll, der im Paket übertragenen Daten
- Nutzdaten Länge (payload length): Größe des AH-Paketes
- reserviert (RESERVED): reserviert für zukünftige Nutzung
- Security Parameters Index (SPI): identifiziert in Verbindung mit der IP-Adresse und dem Sicherheitsprotokoll die Sicherheitsassoziation
- Feld mit Sequenznummern (sequence number): ansteigende Nummer, die vom Absender gesetzt wird, soll Schutz vor Replay-Angriff bieten
- Authentizitätsdaten (authentication data): enthält den Wert des Integritätstests (integrity check value, ICV)

Encapsulated Security Payload (ESP)

Encapsulating Security Payload (ESP) soll die Authentifizierung, Integrität und Vertraulichkeit von IP-Paketen sicher stellen. Im Unterschied zu AH wird der Kopf des IP-Paketes nicht mit berücksichtigt.

Ein ESP-Paket sieht folgendermassen aus:



Bedeutung der Felder

- Security Parameters Index (SPI): identifiziert in Verbindung mit der IP-Adresse und dem Sicherheitsprotokoll die Sicherheitsassoziation
- Sequenznummern (sequence number): ansteigende Nummer, die vom Absender gesetzt wird, soll Schutz vor Replay-Angriff bieten
- Nutzdaten (payload data): enthält die Datenpakete
- Füllung (padding): wird für eingesetzte Blockchiffre genutzt, um Daten bis zur vollen Größe des Blocks aufzufüllen
- Länge Füllung (pad length): enthält Anzahl der eingefügten Bits für Padding
- Nächster Header (next header): identifiziert das Protokoll, der im Paket übertragenen Daten
- Authentizitätsdaten (authentication data): enthält den Wert des Integritätstests (integrity check value, ICV)

Kritik an IPsec

IPsec was a great disappointment to us. Given the quality of the people that worked on it and the time that was spent on it, we expected a much better result. (Bruce Schneier)

Die Experten für Kryptographie Bruce Schneier und Niels Ferguson evaluierten mehrfach das IPsec-Protokoll und fanden mehrere Kritikpunkte. Neben der Art wie es entstand, wird vor allem die hohe Komplexität und damit Fehleranfälligkeit kritisiert. Allerdings stellen beide auch fest, dass es das beste derzeit erhältliche IP-Sicherheits-Protokoll ist.

Beim Lesen der RfCs stellt man fest, dass es keinerlei Angaben gibt, was man mit IPsec erreichen wollte. Weiterhin gibt es in den RfCs z.T. widersprüchliche Angaben.

Wenn man sich beide Modi *Tunnel* und *Transport* betrachtet, stellt man fest, dass der Transportmodus eine Teilmenge des Tunnelmodus ist. Mit kleinen Erweiterungen könnte man mit dem Tunnelmodus alles abdecken.

Die Ergebnisse sind unter http://www.schneier.com/paper-ipsec.html zu finden

Security through Obscurity

Security through obscurity oder **security by obscurity** (engl. „Sicherheit durch Unklarheit“) bezeichnet ein Prinzip in der Computer- und Netzwerksicherheit, nach dem versucht wird, Sicherheit durch Geheimhaltung zu erreichen.

Wenn zum Beispiel jemand den Schlüssel seiner Haustüre in einem Blumentopf versteckt, für den Fall, dass er sich aus dem Haus ausschließt, verlässt er sich auf **Security through obscurity**. Der theoretische Schwachpunkt dieser Vorgehensweise ist offensichtlich: Jeder, der weiß, wo der Schlüssel versteckt ist, kann die Haustüre öffnen. Der Hauseigentümer nimmt jedoch an, dass niemand von dem Versteck weiß und auch ein Einbrecher den Schlüssel kaum finden würde.

The enemy knows the system. – Claude Shannon

„Der Feind kennt das System“, das ist der Ansatzpunkt, von dem man bei der Erstellung von Sicherheitskonzepten ausgehen sollte. Sicherheit, die *nur* auf der Geheimhaltung von Informationen beruht, stellt sich sehr oft als ungenügend heraus.

Beispiele

Häufig wird security through obscurity von unwissenden Benutzern dilettantisch angewendet:

Portscans „ignorieren“: Software wie Personal Firewalls melden Portscans fälschlicherweise oft als „Angriff“ auf den Computer. Viele Benutzer konfigurieren ihre Personal Firewall deshalb so, dass Anfragen auf Ports ignoriert (*DENY*) anstatt beantwortet werden (*REJECT*) und hoffen so, *unsichtbar* und sicherer zu sein. In diesem Fall werden sogar legitime Benutzer und Server behindert.

Ping „ignorieren“: Einige Hosts sind aus den gleichen Gründen wie bei Portscans so konfiguriert, dass sie *ICMP Echo Request*-Pakete ignorieren. Das erhoffte Unsichtbar Sein ist allerdings ebenfalls ein Trugschluss. Ist man nämlich tatsächlich nicht online, bekommt man von einem Router ein IC*MP Destination unreachable* als Antwort. Das heißt, bekommt man *überhaupt keine* Antwort, weiß man, dass der Computer online ist, aber nicht antwortet.

Ernstere Beispiele in der Netzwerksicherheit:

IP-Adressen „verbergen“: Mit NAT oder Masquerading lässt sich die interne Netzwerkstruktur nach außen hin verbergen.

Closed Source-Software: Wie sich Open Source und Closed Source unter dem Aspekt der Sicherheit verhalten, ist sehr umstritten. Linux zum Beispiel profitiert davon, dass der Quelltext von vielen Programmierern durchgesehen wird und so auch Programmfehler gefunden werden.

In diesem Zusammenhang wird oft Linus' Law zitiert (ursprünglich von Eric Raymond):

> *Given enough eyeballs, all bugs are shallow.*

Passwörter: Das sehr weit verbreitete Konzept von *Passwörtern* ist auf den ersten Blick auch security through obscurity: Man hält ein Passwort geheim, um sicher zu gehen, dass nur Befugte Zugang oder Zugriff auf Etwas haben.

Dieses Konzept besteht aus den zwei Teilen Passwort und (Passwort-)Eingabemaske, die einander bedingen. Wenn man annimmt, dass man *beide* Komponenten *tatsächlich* geheimhalten könnte, wäre das Passwort-Konzept ad absurdum geführt. Man spricht deshalb nur dann von security through obscurity, wenn versucht wird, beide Komponenten geheimzuhalten.

Gefahren im Internet

Cracker

Cracker sind Personen mit Computerfachkenntnissen, die im Gegensatz zu Hackern ihre Fähigkeiten grundsätzlich destruktiv einsetzen. Dazu gehört das mutwillige oder kriminelle Eindringen in fremde Computersysteme, auch mit Übernahme der Kontrolle über das fremde System, Diebstahl von Rechenleistung für eigene Zwecke oder Diebstahl, Manipulation oder Zerstörung von Daten.

Der Ursprung des Begriffs liegt in der englischen Umgangssprache bzw. Slang und bezeichnet hier das Aufbrechen von etwas oder das (Zer-) Brechen der Wirkung eines Sicherheitssystems oder einer Sperrvorrichtung.

Der Begriff **Cracker** bezeichnet ebenfalls eine Person, die widerrechtlich Schutzmechanismen (Kopierschutz) kommerzieller Software durch Manipulation des Quell- oder Binärcodes zum Zwecke der illegalen Benutzung oder Verbreitung (Software-Piraterie) außer Kraft setzt. Software-Cracker sind meist in Gruppen streng organisiert. Beispiele für solche Gruppen sind z. B. BLiZZARD, CORE und FairLight, wobei die ersteren sich um Anwendersoftware „kümmern“, letztere um Spiele. Verschiedene meist jugendliche Software-Cracker sind dabei nicht an Profit interessiert, sondern betrachten das Cracken von Spielen gegen die Zeit als Wettbewerb der Gruppen gegeneinander.

Seit dem kommerziellen Erfolg des World Wide Web treten vermehrt geltungssüchtige Teenager mit nur geringen Computer-Kenntnissen als Cracker im Systembereich auf. Diese suchen im Internet nach konfigurierbaren Cracker-Programmen, die sie gegen fremde Systeme einsetzen, um sich in der Gruppe zu profilieren. Leider gibt es noch kein Gesetz zum Verbot gegen das Anbieten von Viren-ähnlichen Programmen im Internet (nur die Anwendung gegen andere ist illegal). Diese Art von Crackern wird von Fachleuten abschätzig als Script Kiddie bezeichnet, da sich ihre Aktivität auf die simple Anwendung von vorgefundenen Schadprogrammen beschränkt, die oft in Script-Sprachen vorliegen, deren Programmierung diese Cracker aber kaum selbst beherrschen.

Von den Behörden und geschädigten Unternehmen aufgespürte und zur Rechenschaft und zum Schadenersatz herangezogene Cracker erlangen teilweise traurige Berühmtheit. Strafen schließen (in den USA) zuweilen mehrjährige Verbote des Umgangs mit Computern jeglicher Art ein.

Literatur

- Clifford Stoll: *Kuckucksei: Die Jagd auf die deutschen Hacker, die das Pentagon knackten.*, Fischer Taschenbücher, ISBN 3-596-13984-8
- *Hacker's Guide*, Markt und Technik, ISBN 3-8272-6522-3

- Markus Schumacher: *Hacker Contest*, Xpert.press, ISBN 3-540-41164-X

Denial of Service

DoS (*Denial of Service*) oder **DDoS** (*Distributed Denial of Service*) sind Angriffe auf Server mit dem Ziel sie bzw. einen oder mehrere ihrer Dienste arbeitsunfähig zu machen. Erfolgt der Angriff von vielen verteilten Systemen aus, wird von einem Distributed Denial of Service Attack (DDoS) gesprochen.

Funktionsweise

Dienste eines Servers, z. B. der HTTP-Server werden mit einer größeren Anzahl Anfragen belastet als dieser in der Lage ist zu bearbeiten, woraufhin er eingestellt wird oder reguläre Anfragen so langsam beantwortet, das diese abgebrochen werden.

Im Unterschied zu anderen Angriffen will der Angreifer hier normalerweise nicht in den Computer eindringen und benötigt deshalb keine Passwörter oder ähnliches. Jedoch kann ein DoS-Angriff Bestandteil eines Angriffs auf ein System sein, z.B. bei folgenden Szenarien:

- Um vom eigentlichen Angriff auf ein System abzulenken, wird ein anderes System durch einen DoS lahmgelegt. Dies soll dafür sorgen, dass das mit der Administration betraute Personal vom eigentlichen Ort des Geschehens abgelenkt ist, bzw. die Angriffsversuche im durch den DoS erhöhten Datenaufkommen untergehen.
- Verzögert man Antworten eines regulären Systems, können Anfragen an dieses durch eigene, gefälschte Antworten kompromittiert werden. Beispiel hierfür ist die "Übernahme" fremder Domainnamen durch Liefern gefälschter DNS-Antworten.
- Als Form des Protests sind DoS-Attacken in letzter Zeit populär geworden. Zum Eigenschutz der Protestierenden werden Angriffe dieser Art im allgemeinen durch Würmer durchgeführt, die sich selbstständig auf fremden Systemen verbreiten. Entsprechend handelt es sich bei Protestaktionen dieser Art um DDoS-Attacken.

Beispiele

- Februar 2004: Der Emailwurm MyDoom bringt die Website der Firma SCO zum Erliegen.
- Mai 2003: Schlechte Programmierung bei den Internet-Routern der Firma Netgear führte dazu, dass sämtliche Geräte weltweit ihre Zeit bei einem Server der Universität von Wisconsin abgleichen wollten. Dies führte dazu, dass die Bandbreite für die Anfragen mehr als 100 MBit/s betrug und der Zusammenbruch des Zeitservers erfolgte.

- August 2003: Der Emailwurm Lovsan/W32.Blaster soll die Update-Site der Firma Microsoft unerreichbar machen, wird jedoch durch Deaktvierung des Domainnamens ins Leere geführt.
- Im Februar 2000 wurden verschiedene, große Internet-Dienste (wie z.B. Yahoo!, CNN, Amazon.de, eBay) durch DDoS-Attacken lahm gelegt. Hierbei hatten sich die Angreifer Zugang zu hunderten von Computern im Internet verschafft (darum das "distributed" also "verteilt"), um die Wirksamkeit ihrer Attacken durch die Vielzahl der gleichzeitig angreifenden Rechner stark zuerhöhen. Eine DDoS-Attacke erzielt den Schaden meistens durch die Überlastung der angegriffenen Systeme.
  Die beobachteten Angriffe basierten auf zwei wesentlichen Schwachstellen:
  Zum einen konnten die Absenderadressen der "angreifenden" Datenpakete gefälscht werden (IP-Spoofing), zum anderen konnte vor den eigentlichen Angriff auf einer großen Anzahl dritter - nur unzureichend geschützter - Internet-Rechner unberechtigterweise Software installiert werden, die dann ferngesteuert durch massenhaft versendete Datenpakete den eigentlichen Angriff ausführten. Das besondere an diesen DDoS-Angriffen ist, dass diese daher auch diejenigen treffen können, die sich ansonsten optimal vor Eindringlingen aus dem Internet geschützt haben. Insofern sind Rechner, auf denen noch nicht einmal so genannte Grundschutzmaßnahmen umgesetzt sind, nicht nur für den jeweiligen Betreiber eine Gefahr, sondern auch für alle anderen Computer im Internet.

GEGENMASSNAHMEN

Wirksame Maßnahmen gegen verteilte Denial-of-Service-Angriffe müssen in einer konzertierten Aktion an vielen Stellen in der vorhandenen komplexen Internetstruktur getroffen werden. Serverbetreiber im Internet, die Ziel der genannten Angriffe waren, können eine Reihe von sinnvollen Maßnahmen ergreifen, aber das DDoS-Problem nicht vollständig lösen. Vielmehr müssen verschiedene Zielgruppen (Inhalte-Anbieter, Serverbetreiber, Netzvermittler und Endanwender) - jeder in seinem Bereich - tätig werden. Nur gemeinsam kann das Internet im Hinblick auf die Gefährdung durch DDoS-Angriffe sicherer gemacht, die Durchführung von Denial-of-Service-Angriffen erschwert sowie eine spätere Verfolgung der Urheber dieser Angriffe erleichtert werden.

Denial-of-Service Attacken sind immer böswillig, außer wenn Sicherheitsexperten DoS-Attacken gegen ihre eigenen Netzwerke ausführen um die Sicherheit ihres eigenen Netzwerkes zu prüfen (so genanntes Auditieren). DoS-Attacken können strafbar sein und eine Strafanzeige nach sich ziehen.

DoS-Angriffe können auf jeder Plattform stattfinden. Eine einzelne DoS-Attacke kann mehrere Zielbetriebssysteme treffen (zum Beispiel konnte die Land-Attacke fast zwei Dutzend verschiedene Betriebssysteme beeinträchtigen, darunter Windows NT und einige UNIX-Versionen).

IP-SPOOFING

IP-Spoofing bezeichnet in Computernetzen das Versenden von IP-Paketen mit gefälschter Quell-IP-Adresse.

Der Header jedes IP-Pakets enthält dessen Quelladresse. Dies sollte die Adresse sein, von der das Paket gesendet wurde. Indem er den Header so fälscht, dass er eine andere Adresse enthält, kann ein Angreifer das Paket so aussehen lassen, als ob das Paket von einer anderen Maschine gesendet wurde. Dies kann von Eindringlingen dazu genutzt werden, Sicherheitsmaßnahmen wie z. B. IP-Adressbasierte Authentifizierung im Netzwerk auszutricksen oder zum Verschleiern des eigenen Rechners dienen.

Diese Art von Angriff ist am effektivsten, wenn zwischen den Maschinen in einem Netzwerk Vertrauensbeziehungen bestehen. In manchen Firmennetzen ist es durchaus üblich, dass interne Systeme sich gegenseitig vertrauen, so dass ein Benutzer sich ohne Benutzernamen und Passwort einloggen kann, wenn er von einer anderen internen Maschine auf das Netzwerk zugreift und daher bereits auf einem anderen Rechner eingeloggt ist. Indem nun eine Verbindung von einer vertrauenswürdigen Maschine gefälscht wird, könnte ein Angreifer den Zielrechner angreifen, ohne sich zu authentifizieren.

GEGENMASSNAHMEN

Paketfilter sind eine mögliche Gegenmaßnahme gegen IP-Spoofing. Das Gateway zu einem Netzwerk sollte eine *eingehende Filterung* vornehmen: Von Außen kommende Pakete, die Quelladressen von innenliegenden Rechnern haben, werden verworfen. Dies verhindert, dass ein externer Angreifer die Adresse einer internen Maschine fälschen kann. Idealerweise sollten auch ausgehende Pakete gefiltert werden, wobei dann Pakete verworfen werden, deren Quelladresse nicht innerhalb des Netzwerks liegt; dies verhindert, dass IPs von externen Maschinen gespooft werden können und ist eine bereits lange bestehende Forderung von Sicherheitsfachleuten gegenüber Internet Service Providern: Wenn jeder ISP konsequent ausgehende Pakete filtern würde, die laut ihrer Quelladresse nicht aus dem eigenen Netz stammen, wäre massenhaftes IP-Spoofing (häufig in Verbindung mit Denial of Service-Attacken) ein wesentlich geringeres Problem als es heute im Internet ist.

Einige Protokolle auf höheren Schichten stellen eigene Maßnahmen gegen IP-Spoofing bereit. Das Transmission Control Protocol (TCP) benutzt beispielsweise Sequenznummern, um sicherzustellen, dass ankommende Pakete auch wirklich Teil einer aufgebauten Verbindung sind. Die schlechte Implementation der TCP-Sequenznummern in vielen älteren Betriebssystemen und Netzwerkgeräten führt jedoch dazu, dass es dem Angreifer unter Umständen möglich ist, die Sequenznummern zu erraten und so den Mechanismus zu überwinden. Alternativ könnte er versuchen, zum *Man in the Middle* zu werden.

Sicherheits-Implikationen

IP-Spoofing lässt sich für sich genommen nur beschränkt zum Einbruch in andere Systeme benutzen, da alle Antwortpakete des angegriffenen Rechners an die gefälschte Adresse gesendet werden. Umgekehrt lässt sich dieses Verhalten jedoch auch als „Waffe“ benutzen, wenn mit gespooften Paketen SYN-Flooding betrieben wird; hierzu sendet man gefälschte Pakete an bestimmte Rechner, und die Antwortpakete landen bei dem als Quelladresse angegebenen Opfer, dessen Verbindung dadurch möglicherweise lahmgelegt wird. Die Identität des tatsächlichen Angreifers ist dabei nur schwer feststellbar, da die Quelle der Antwortpakete natürlich der vorher überrumpelte arglose Rechner ist.

Protokoll-Spoofing wird auch zur Datenkompression verwendet, und wurde erstmals 1985 vom Hayes Smartmodem benutzt, das Teile des UUCP-Protokolls spoofte, um den Datendurchsatz zu steigern. Dabei wurden Protokollheader gestrippt oder komplett entfernt, und auf der anderen Seite wieder rekonstruiert.

Computervirus

In der Fachsprache ist ein **Computervirus** eine nicht selbständige Programmroutine, die sich selbst reproduziert, indem sie sich an andere Computerprogramme oder Bereiche des Betriebssystems anhängt und, einmal gestartet, vom Anwender nicht kontrollierbare Manipulationen an selbigen vornimmt.

Umgangssprachlich hat der Begriff **Computervirus** eine breitere Bedeutung: Er wird sowohl für Computerviren in der fachsprachlichen Bedeutung als auch für Computerwürmer, Trojanische Pferde benutzt.

Die Eigenschaft der Reproduktion führte wie beim biologischen Vorbild zu der Bezeichnung „Virus“.

Durch Computerviren kommt es auf einem Computer häufig zu Verfälschung oder Verlust von Daten und Programmen, sowie zu Störungen des regulären Betriebs.

Aufbau

Es gibt mehrere Versuche, einen Virus zu strukturieren:

Erklärung 1

- **Vermehrungsteil:** Mit diesem Programmteil wird die Vermehrung des Virus durchgeführt.
- **Erkennungsteil:** Im Erkennungsteil wird geprüft, ob bereits die Infektion eines Programms oder Systembereichs erfolgte. Jedes Wirts-Programm wird nur einmal infiziert.

- **Schadensteil:** In einigen Viren ist absichtlich eine Schadensfunktion programmiert, meist das Überschreiben oder Verändern von Programmen oder Daten, oder aber auch nur die Ausgabe von Meldungen und Geräuschen auf dem Rechner. Dieser Programmteil kann fehlen, aber auf jeden Fall entsteht Schaden durch Inanspruchnahme von Speicherplatz im Hauptspeicher und auf Datenträgern. Durch Programmierfehler, Veränderungen des Betriebssystems oder ähnliches können weitere Schäden als Nebeneffekte auch dann auftreten, wenn sie nicht absichtlich programmiert sind.
- **Bedingungsteil:** Sowohl die Verbreitung als auch die Schadensfunktion können von Bedingungen abhängig programmiert sein, z. B. tritt bei einigen Viren der Schaden an einem bestimmten Datum oder bei einer bestimmten Anzahl von Aufrufen ein. Auch dieser Teil kann fehlen.
- **Tarnungsteil:** Hierunter fallen Programmroutinen, um die Entdeckung des Virus im infizierten System zu erschweren. Dieser Teil ist meist nur bei neueren Viren zu finden.

Erklärung 2

Ein Computervirus besteht aus drei Teilen:

- Replikationseinheit
- Trigger
- Payload (Teil, der den Schaden anrichtet)

Die **Replikationseinheit** dient dazu, dass der Virus sich unbemerkt vervielfältigen kann, z. B. indem er sich an ein bestimmtes Programm anhängt, welches sich auf der Festplatte befindet. Jedesmal, wenn das infizierte Programm gestartet wird, kopiert sich der Virus. Die Replikationseinheit dient also dazu, dass der Virus sich erst einmal verbreiten kann, ohne zunächt einen Schaden anzurichten. Andere Verbreitungsmethoden sind die Infektion von Bootdisketten (MBR Viren, heute Irrelevant) oder Text Dokumenten (falls das Textverarbeitungsprogramm eine genügend mächtige Makro Sprache besitzt, und die Makros zusammen mit dem normalen Text in einder Datei gespeichert werden).

Der Teil, der den eigentlichen Schaden verursacht, bezeichnet man als **Payload** (dtsch. Nutzlast). Hier tritt der Virus zum ersten mal mit einer Meldung in Erscheinung und signalisiert, dass er da ist. Kurz darauf richtet er einen mehr oder weniger großen Schaden an. Der Payload ist optional (ein Virus ist auch ohne ihn ein Virus) und in den meisten Viren gar nicht vorhanden.

Einige Viren sind so programmiert, dass sie erst dann in Erscheinung treten und den Schaden verursachen, wenn ein bestimmtes Ereignis eingetreten ist oder eine bestimmte Zeit verstrichen ist. Andere schalten gewisse Funktionen zu einem bestimmten Zeitpunkt wieder ab. Hierfür ist der sogenannte **Trigger** zuständig.

So kann es z. B. sein, dass der Virus erst nach dem 100sten Start des Rechners aktiv wird, oder dass er sich jeden Dienstag oder am 10ten eines jeden Monats meldet und dann z. B. bestimmte Dateien löscht.

Grundtypen sind Boot-Viren, Datei-Viren und Makro-Viren

Geschichte

- 1981 - Professor Leonard M. Adleman verwendet im Gespräch mit Fred Cohen zum ersten Mal den Begriff *Computervirus*
- 1984 - Fred Cohen liefert seine Doktorarbeit „Computer Viruses - Theory and Experiments“ ab. Darin wurde ein funktionierender Virus für das Betriebssystem UNIX vorgestellt.
- 1985 - Über Mailboxen wird das Trojanische Pferd Gotcha über ein Programm verteilt das die Grafik verbessern soll. Nach dem Start werden die Daten auf der Festplatte gelöscht und es erscheint auf dem Bildschirm der Schriftzug *Arf, arf, Gotcha*.
- 1986 - Zwei Software-Händler aus Pakistan verbreiten den ersten Virus für das Betriebssystem DOS. Das Programm war relativ harmlos da es nur das Inhaltsverzeichnis der befallenen Disketten in *Brain* umbenannte.
- 1987 - Der so genannte *Cascade*-Virus läßt zum ersten Mal in Deutschland die Buchstaben einer Seite nach unten rutschen wo sie sich zu einem kleinen Häufchen sammeln. Er vernichtete Dateien.
- 1988 - Der erste Baukasten für Viren (Virus Construction Kit) wird veröffentlicht. Damit ist es auch Anfängern möglich Viren nach Maß zu erstellen. Das Programm läuft auf dem Computer Atari ST.
- 1989 - Mit V2Px erscheint der erste polymorphe Virus der sich selbst wieder neu verschlüsseln kann und deshalb durch Anti-Virus-Software nur schwer zu entdecken ist.
- 1990 - Der *Verband deutscher Virenliebhaber* verbreitet das erste Virus Construction Kit für DOS.
- 1993 - Erste Computerviren für Windows tauchen auf.
- 1995 - Es erscheinen die ersten Makroviren.
- 1997 - Der erste Virus für das Betriebssystem Linux taucht auf.
- 1998 - *Strange Brew* der erste Virus für Java erscheint.

Prävention

Viren werden entweder vom Anwender selbst (oft unabsichtlich) oder von unsicherer Software installiert. Anwender sollten deshalb niemals unbekannte Programme ausführen. Vor allem Microsoft Outlook und Outlook Express sind als sehr unsichere Mail-Clients aufgefallen, da sie ohne Zutun des Benutzers fremde Software in E-Mails gestartet haben. Man sollte deshalb überlegen, ein sichereres Programm zu benutzen.

Antivirenprogramme schützen nur vor bekannten Viren. Unbekannte Viren können jedoch von manchen dieser Programme auch anhand ihres Verhaltens entdeckt werden. Diese Funktionen arbeiten jedoch extrem unzuverlässig. Aus diesen Gründen sollte man diese Programme nur als Unterstützung ansehen und sich nicht allein auf ihr Urteil verlassen

Personal Firewalls können theoretisch vor bösartiger Software, die sich über Schwachstellen in Serverdiensten weiterverbreitet, schützen. In der Praxis ist es jedoch besser, die kritschen Dienste zu beenden, da jedes Programm mit Internetzugriff, auch eine Personal Firewall, ein potentielles Angriffsziel darstellt. Des Weiteren sind Personal Firewalls gegen Computerviren fast immer unwirksam, da diese sich im Allgemeinen durch die Weitergabe infizierter Dateien durch die Benutzer verbreiten.

Folgende Tipps helfen, die Bedrohung durch Viren einzuschränken:

1. Dateien aus dem Internet (ob nun per Download heruntergeladen oder per E-Mail erhalten) sollten nur angenommen werden, wenn man sicher ist, dass sie aus seriöser Quelle stammen (E-Mail-Absender können gefälscht sein)
2. Das automatische Öffnen von Dateien aus dem Internet sowie das automatische Anzeigen von Dateianhängen sollte deaktiviert werden.
3. Regelmäßig Betriebssystem und Software aktualisieren
4. Einen sicheren Browser und ein sicheres E-Mail-Programm verwenden

eventuell auch:

1. Kein Windows verwenden sondern MacOS, Linux oder ein Unix-Derivat; diese Systeme sind zwar nicht unangreifbar, aber es gibt weniger Personen, die maliziöse Software für diese herstellen und verbreiten. Auch ist es auf diesen Systemen einfacher, konsequent den Benutzer vom Systemverwalter zu trennen.
2. Verwendung aktueller Antivirenprogramme mit Virendefinitionen, die mindestens wöchentlich aktualisiert werden.
3. Einsatz einer Personal Firewall, um Angriffe abzublocken, die Sicherheitslücken im System ausnutzen.

Computerwurm

Ein **Computerwurm** ist eine Programmroutine, die sich selbst reproduziert, indem sie über ein Computernetzwerk an Computerprogrammen oder Betriebssystemen anderer Computern Manipulationen vornimmt.

Ein Wurm kann eine spezielle Schadensroutine enthalten, muss aber nicht. Da ein Wurmprogramm auf befallenen Systemen Ressourcen zur Weiterverbreitung bindet, können selbst Würmer ohne spezielle Schadensroutinen gewaltige wirtschaftliche Schäden erzeugen.

Viren und Würmer

Würmer sind den Computerviren konzeptionell sehr ähnlich. Die Abgrenzung besteht darin, dass ein Virus versucht, Dateien auf einem Computersystem zu infizieren, während ein Wurm versucht, eine Zahl von Computern in einem Netzwerk zu infizieren. Außerdem benötigt ein Virus ein Wirtsprogramm, welches er infiziert. Wird dieses Programm ausgeführt

wird gleichzeitig auch der Virus ausgeführt. Bei einem Wurm handelt es sich dagegen um ein eigenständiges Programm.

Tarnung und Verbreitung

Heutzutage verbreiten sich Würmer überwiegend per E-Mail, wobei sie sich als Datei-Anhang an einen kurzen Text anfügen. Verschiedene Mechanismen dienen zum Tarnen des gefährlichen Anhangs bzw. zum automatischen Ausführen:

Tarnung

Würmer bekommen Dateinamen mit doppelten Erweiterungen, z. B. music.mp3.exe, wobei darauf gebaut wird, dass der Anwender die Optionen für die Dateiansicht von Windows nicht geändert hat. Diese werden bei der Installation von Windows so eingestellt, dass Endungen von bekannten Anwendungen nicht angezeigt werden. Der Empfänger des genannten Anhangs sieht also nur music.mp3 und klickt auf den Anhang, um die vermeintliche Musik abzuspielen - schon ist der Virus aktiviert. Ein Beispiel:

```
Content-Type: audio/mpeg;
name="music.mp3"
Content-Transfer-Encoding:base64
Content-Disposition: attachment;
filename="music.mp3.exe"
```

Oft werden Würmer in ZIP-Archive verpackt, um es Virenscannern zu erschweren, den Schädling zu analysieren.

Automatisches ausführen

Zum automatischen ausführen enthält die Email HTML-Code der ein Fenster im Fenster (iframe) erzeugt, in dem der Datei-Anhang mit Hilfe von Javascript gestartet wird. Der Wurm verschickt sich selbst, wobei aus der Adressbuch des Benutzers wahllos Absenderadressen entnommen werden. Es ist daher relativ sinnlos, beim Empfang einer verseuchten Email eine Warnung an die Absenderadresse zu schicken, es trifft höchstwahrscheinlich den Falschen.

Geschichte

Das Konzept eines Computerwurms oder Netzwerkwurms wurde erstmals 1975 im Science-Fiction Buch „**The Shockwave Rider“** von John Brunner erwähnt.

Robert Morris programmierte 1988 den Morris-Wurm, der zwar keine Schadensroutine enthielt, aber durch seine aggressive Weiterverbreitung unter Ausnutzung von einigen Unix-Diensten, wie z. B. Sendmail, fingerd oder rexec sowie der r-Protokolle ca. 6000 Rechner, das entsprach damals ungefähr 10% des weltweiten Netzes, lahm legte. Allerdings hatte das Morris nicht beabsichtigt - es war ein Programmierfehler, der zu diesem Debakel führte.

Im März 1999 verbreite sich, über Outlook der E-Mail-Wurm Melissa. Er führte allerdings keine zerstörende Aktion aus, sondern löste nur eine E-Mail-Flut aus.

Ins öffentliche Bewusstsein gerieten Würmer spätestens 2000 mit dem massiven Auftreten des ILOVEYOU E-Mail Wurms, der viele Nachahmer inspiriert hat.

Neben den E-Mail Würmern ist heute die beliebteste Weiterverbreitungsstrategie das Ausnutzen von Sicherheitslücken in häufig installierter Software. Prominente Beispiele sind „Code Red“, der 2001 in einer großen Welle Systeme mit dem Microsoft Internet Information Server attackierte, und „SQL Slammer“ der 2003 eine Sicherheitslücke im Microsoft SQL Server ausnutzte als Verbreitungsstrategie.

Mitte 2003 machte der Wurm W32.Blaster Schlagzeilen. Er verbreitete sich aufgrund einer Sicherheitslücke im Microsoft Windows Betriebssystem, um einen Distributed Denial of Service-Angriff auf Microsoft Server vorzubereiten. Etwa zur gleichen Zeit befiel der Wurm Sobig.F mit massiver Geschwindigkeit unzählige Computer und belastete Mailserver auf der ganzen Welt.

Am 26. Januar 2004 wurde der Wurm Mydoom das erste Mal gesichtet. Die schnelle Verbreitung des Wurms sorgte für ein paar Stunden zu einer durchschnittlich 10-prozentigen Verlangsamung des Internetverkehrs und einer durchschnittlich erhöhten Ladezeit der Webseiten von 50 Prozent.

Am 3. März 2004 machten neue Varianten des Mitte Februar im Internet aufgetauchten Bagle-Wurms Schlagzeilen. Sie können den zentralen Virenschutz in Firmennetzwerken unterlaufen.

April/Mai 2004 taucht mit Phatbot ein neuer Wurm auf, der mit anderen Würmern wie Sasser, Beagle und Mydoom interagiert.

Trojanisches Pferd

Als **Trojanisches Pferd** oder umgangssprachlich auch **Trojaner** (engl. *Trojan* - fälschlicherweise, weil die *Trojaner* ja selbst die Opfer des *Trojanischen Pferdes* geworden sind) bezeichnet man in der Computersprache auch ein Fernwartungsprogramm, das speziell für bösartige Zwecke eingesetzt wird.

Im einfachsten Fall handelt es sich bei einem Trojanischen Pferd um ein Computerprogramm, das etwas völlig anderes tut, als seine Beschreibung verspricht. Das Trojanische Pferd kann „huckepack“ mit einem anderen Programm auf den Rechner des Anwenders gelangen und sich im Schatten des eigentlich gewünschten Programms installieren.

Wenn ein Trojanisches Pferd installiert ist, läuft es als Serverprozess auf dem Computer. Das heißt andere Netzwerk- oder sogar Internetbenutzer können sich mit dem Trojanischen Pferd verbinden und direkt auf dem fremden Computer Veränderungen durchführen oder Daten ausspähen. So kann das Trojanische Pferd etwa sensible Daten (gespeicherte Passwörter, Kreditkartennummern, Kontonummern und ähnliches) ausspähen oder unbemerkt Dateien versenden.

Trojanische Pferde werden auch häufig als Updates von bekannten Computerprogrammen angeboten, oft im Rahmen von Tool-Sammlungen im World Wide Web, die versprechen, mehr zu können als vergleichbare Programme von bekannten Firmen.

Von einem Computervirus unterscheidet sich ein Trojanisches Pferd dadurch, dass es sich nicht selbst reproduzieren kann, sondern darauf angewiesen ist, von einem Anwender kopiert zu werden: per Download, als Dateianhang einer E-Mail oder durch Weitergabe an andere Benutzer.

Am besten schützt man sich vor Trojanischen Pferden, indem man keine Programme aus unbekannten Quellen ausführt. Wie auch bei Viren, schützen Antivirenprogramme, wenn überhaupt (einige Antivirenprogramme schützen gar nicht vor Trojanischen Pferden), nur vor bekannten Trojanischen Pferden. Personal Firewalls bieten keinen Schutz vor der Installation dieser Software. Allerdings kann man auf die eventuelle Kommunikation des Trojanischen Pferdes durch eine Firewall aufmerksam werden oder Sie durch geeignete Firewallregeln gar nicht erst zulassen.

Backdoor

Als **Backdoor** (deutsch **Hintertür**) bezeichnet man in einem Computersystem eine Funktionalität, die es ermöglicht, unter Umgehung der normalen Zugriffssicherung Zugang zum Computer (oder einem Computerprogramm) zu ermöglichen. Dies geschieht üblicherweise derart, dass der Zugriff den Administratoren verborgen bleibt.

Eine Variante sind in einem System fest vorgegebene, nur dem Ersteller des Systems bekannte Passwörter oder andere versteckte Funktionen, die den Login ohne die sonst übliche Authentifizierung ermöglichen. Eines der Argumente für Open-Source-Software ist es, dass der Quelltext nach derartigen Hintertüren leicht von jedem selbst durchsucht werden kann. Im Gegensatz dazu seien proprietäre Anwendungen nicht in ihrer Funktionalität einsehbar.

Eine harmlose Variante dieser Art der Manipulation wird als Osterei bezeichnet, bei der über bestimmte Eingabemethoden versteckte Programmteile aufgerufen werden können (oft erscheinen die Namen der Programmierer oder ein kleines Spiel).

Bösartige Hintertüren werden häufig von Computerviren, Würmern oder Hackern auf einem kompromittierten System installiert, nachdem beispielsweise über eine Sicherheitslücke in das System eingedrungen wurde. Diese so mit einer Hintertür versehenen Computer dienen dann meist dazu, Ausgangspunkt für Angriffe auf weitere Systeme zu sein, da so die Herkunft des Angriffs über mehrere Systeme hinweg verschleiert werden kann.

Rootkit

Ein **Rootkit** ist eine Sammlung von Softwarewerkzeugen, die nach dem Hack eines Computersystems auf dem kompromittierten System installiert wird, um zukünftige Logins des Hackers zu verbergen, Prozesse zu verstecken und Daten mitzuschneiden.

Der Name Rootkit entstand aus der Tatsache, dass die ersten Sammlungen von Unix-Tools zu oben genannten Zwecken aus modifizierten Versionen der Programme ps, netstat, passwd usw. bestanden, die dann jede Spur des Hackers, die sie normalerweise zeigen würden, verbargen, und es dem Hacker so ermöglichten, mit den Rechten des Systemadministrators root zu agieren, ohne dass der wirkliche Administrator dies bemerken konnte.

Rootkits, die modifizierte Programme einsetzen, um sich selbst zu verbergen, sind jedoch relativ einfach durch den Vergleich der Prüfsummen der Programmdateien aufzuspüren. Die **LKM-Rootkits** hingegen verbergen sich, indem sie spezielle Programmteile in den Betriebssystem-Kernel als nachladbare Kernel-Module installieren (LKM steht für engl. *loadable kernel module*), und einige Funktionen des Betriebssystems ersetzen.

Der Begriff ist heute nicht mehr allein auf Unix-basierte Betriebssysteme beschränkt, da es inzwischen Tools gibt, die ähnliche Funktionalität auch für Nicht-Unix-Systeme bieten, obwohl diese natürlich keinen root-Account haben.

DIALER

Dialer oder zu deutsch: **Einwahlprogramme** sind im engeren Sinne Computerprogramme, mit deren Hilfe über das analoge Telefon- oder das ISDN-Netz eine Verbindung zum Internet oder anderen Computernetzwerken aufgebaut werden kann. So ist bei vielen Betriebssystemen bereits ein Standard-Einwahlprogramm für Verbindungen nach dem Point-to-Point Protocol (PPP) mitgeliefert. Bei Windows nennt sich dieser „DFÜ-Netzwerk“. Das Einwahlprogramm muss gestartet werden, wenn man eine Internet-Verbindung aufbauen möchte, und so lange laufen, bis man die Verbindung nicht mehr benötigt und diese schließt.

Viele Provider bieten Installations-CDs an, die es unerfahrenen Kunden vereinfachen sollen, einen passenden Internetzugang einzurichten. Dies geschieht entweder dadurch, dass ein Eintrag im DFÜ-Netzwerk des Windows-Betriebssystems erstellt wird, oder aber dadurch, dass ein firmenspezifisches Einwahlprogramm (z. B. die AOL-Software) installiert wird. Oft wird dabei im weiteren Sinne nicht nur das Einwahlprogramm selbst, sondern auch dessen Installationsprogramm als „Dialer“ bezeichnet.

Bei einem 0190-Dialer handelt es sich um einen Dialer, der eine Verbindung zu einer Rufnummer mit 0190-Vorwahl herstellt bzw. einrichtet. Gedacht waren sie als einfache, anonyme elektronische Zahlungsmöglichkeit für kostenpflichtige Inhalte, die erst dann verfügbar werden, wenn die Einwahl über die spezielle Rufnummer erfolgt. Heute denkt man jedoch beim Begriff „0190-Dialer“ gewöhnlich an solche Dialer, die von unseriösen, teilweise sogar kriminellen Anbietern verbreitet werden, um schnell viel Geld zu machen.

Mit ähnlichen Tricks wie Viren und Würmer werden die Programme vorwiegend auf PCs mit dem Betriebssystem Windows installiert. Danach baut diese Software – meist ohne das Wissen des Benutzers – neue kostenpflichtige Verbindungen auf, oft zu teuren 0190er-Nummern.

Ein neues (Anfang 2003) Visual Basic-Script installiert zum Beispiel ein Trojanisches Pferd, welches Werte in der Windows-Registry und die Sicherheitseinstellungen des Internet

Explorer verändert, damit ActiveX-Steuerelemente ohne Warnung aus dem Internet geladen werden können. Durch den Aufruf einer solchen Seite oder E-Mail wird der Dialer aus dem Internet heruntergeladen. Das Script schaltet auch den Modemlautsprecher ab und unterdrückt die Meldungen während des Aufbaus einer DFÜ-Verbindung. Davon sind besonders Benutzer der Programme Outlook, Outlook Express und des Internet Explorers betroffen, wenn die Ausführung von ActiveX-Objekten oder JavaScript in den Sicherheitseinstellungen erlaubt ist und die neuesten Sicherheitspatches von Microsoft nicht eingespielt sind.

Benutzer, die sich über DSL mit dem Internet verbinden, sind nicht von Dialern betroffen. Ein Dialer kann zwar heruntergeladen werden, aber eine Einwahl ist über DSL nicht möglich, da es im DSL-Netz keine herkömmlichen Telefonnummern gibt. Falls man aber zusätzlich noch einen ISDN-Adapter oder ein Modem angeschlossen hat, besteht trotzdem die Gefahr, dass der Dialer sich einwählt.

Dubiose Dialer erkennt man an folgenden Merkmalen:

- Beim Anklicken einer Webseite öffnet sich ein Download-Popup.
- Auf der Webseite findet man allenfalls einen versteckten Hinweis auf die entstehenden hohen Kosten.
- Der Download findet auch dann statt, wenn man auf „Abbrechen“ geklickt hat.
- Der Dialer installiert sich automatisch selbst als Standardverbindung, ohne dass es einen Hinweis darauf gibt.
- Der Dialer baut selbstständig unerwünschte Verbindungen auf.
- Der Dialer weist vor der Einwahl nicht auf den hohen Preis der Verbindung hin.
- Der anfallende hohe Preis wird während der Verbindung nicht angezeigt.
- Der Dialer lässt sich gar nicht oder erst mit erheblichem Aufwand wieder deinstallieren.

In jüngerer Zeit werden dubiose Dialer mit Hilfe angeblicher Virenschutzprogramme bei ahnungslosen Internetnutzern installiert: Werbe-Mails von einem angeblichen „AntiVirus Team“ enthalten z. T. im Betreff den Zusatz „Weiterleiten“, bewerben aber per Download-Link ein Programm namens 'downloadtool.exe' oder 'antivirus.exe', das in Wirklichkeit einen 0190-Dialer darstellt. Eine andere Masche sind E-Mails, in denen dem Empfänger für seine Hilfe und Unterstützung gedankt wird und er per Klick einen Blick auf die neue Webseite werfen soll. Wer seine Neugier nicht zügeln kann, auf den wartet dann ein Dialer-Download. Weiter gibt es Grußkarten-Mails, in denen ein Link angegeben ist, der eine Webseite öffnet, auf der den Nutzern des Internet Explorers ein ActiveX-Plug-In aufgenötigt wird, das klammheimlich einen Dialer installiert.

Man sollte daher Links in Werbe-Mails oder Links auf den von Werbe-Mails beworbenen Webseiten niemals anklicken. Auch sollte man einen automatisch gestarteten Download sofort abbrechen, wenn man ihn bemerkt hat. Um sich zu schützen, kann man auch bei seiner Telefongesellschaft eine Sperrung aller 0190-Nummern für den eigenen Anschluss beantragen. Die einmalige Einrichtungsgebühr dafür beträgt 7,50 EURO. Diese Sperrung betrifft dann allerdings auch den Faxabruf von Informationen – die etwa in TV-Sendungen angeboten werden – und gilt auch für Support-Rufnummern.

Im März 2003 sind die ersten (bislang „gutartigen“) Handydialer für WAP-Handys aufgetaucht. Sie basieren auf SMS mit ausführbarem Code.

Gesetzliche Regelungen

Seit dem 15. August 2003 ist in Deutschland das „Gesetz zur Bekämpfung des Missbrauchs von (0)190er/(0)900er Mehrwertdiensterufnummern“ in Kraft getreten.

Dieses Gesetz beinhaltet folgende Punkte:

- Preisangabepflicht der Anbieter
- Preisobergrenzen, Legitimationsverfahren und automatische Trennung
- Registrierung von Anwählprogrammen (Dialer)
- Sperrung von Dialern
- Auskunftsanspruch des Verbrauchers gegenüber der RegTP

Am 4. März 2004 entschied der Bundesgerichtshof in Karlsruhe, dass für Dialernutzung anfallende Gebühren nicht gezahlt werden müssen, wenn der Dialer unwissentlich benutzt wurde (Aktenzeichen III ZR 96/03).

Weblinks

- www.regtp.de (*http://www.regtp.de/*) – Regulierungsbehörde für Telekommunikation und Post
- www.dialerschutz.de (*http://www.dialerschutz.de/*) – Infoseite rund um den Schutz vor Dialern

Spyware

Als **Spyware** wird üblicherweise Software bezeichnet, die persönliche Daten des Benutzers ohne dessen Wissen oder gar Zustimmung an den Hersteller der Software oder an Dritte sendet. Oft wird Spyware verwendet, um Produkte scheinbar kostenlos anzubieten.

Meist dienen die Spyware-Programme dazu, das Surf-Verhalten im Internet zu analysieren, um gezielt Werbebanner oder Popups einzublenden, die auf die Interessen des Benutzers angepasst sind. Die Firmen erhoffen sich daraus eine Steigerung in der Wirksamkeit dieser Werbemethoden.

Zur Entfernung von Spyware gibt es spezielle Software.

Web-Bug

Als **Web-Bugs** (deutsch Web-Wanze) bezeichnet man kleine Grafiken in HTML-E-Mails oder auf Webseiten, die eine Logfile-Aufzeichnung ermöglichen.

Die meist nur 1 mal 1 Pixel großen Bilder sind häufig auch transparent, damit sie nicht auffallen. Wird ein Dokument geöffnet, dann muss diese kleine Bild von einem Server im Internet geladen werden, wobei dieser Download dort registriert wird. So kann der Betreiber des Servers sehen, wann und wie viele Nutzer diesen Web-Bug brauchten, bzw. wann bzw. ob eine E-Mail geöffnet oder eine Webseite besucht wurde.

Private Betreiber einer Webseite können durch das Einbinden eines kostenlosen Web-Bugs, ohne Zugriff auf die Logfiles des Servers Informationen über die Besucher erhalten. So werden auf zahlreichen Internetseiten Zähler (Counter) verwendet, die auf dem selben Prinzip aufbauen, aber bei jedem neuen Besucher ein neues „Bild“, d. h. eine um eins erhöhte Zahl darbieten.

Versender von Massenemails und Spams können durch Einbau eines Web-Bugs in die E-Mail ermitteln

- ob eine E-Mail-Adresse gültig ist
- dass und wann die E-Mail gelesen wurde
- welchen Browser und welches Betriebssystem Sie verwenden.
- welche IP-Adresse, damit Ihren ISP-Provider und möglicherweise sogar ihren Wohnort.

Gegenmittel

- Wird eine E-Mail oder Webseite offline gelesen, kann die Grafik des Web-Bugs nicht vom Server geladen werden und dort also auch nicht registriert werden.
- Man kann auch einfach ein Mailprogramm benutzen, das kein HTML-Emails unterstützt bzw. sie nicht anzeigt.
- Man installiert sich ein Programm gegen Web-Bugs. Diese ersetzen 1x1 Grafiken durch eigene Bilder und verhindern das Laden der Datei von einem Server. (z. B. *http://www.bugnosis.org/*)

Spam

Spam bezeichnet unerwünschte Werbe-E-Mails, die meist in großen Mengen verschickt werden.

Der Begriff entstammt einem Sketch der englischen Komikergruppe Monty Python: In einem Restaurant besteht die Speisekarte ausschließlich aus Gerichten, die Spam enthalten (*SPAM* ist ein Markenname für Dosenfleisch, 1936 entstanden aus **sp***iced h***am**, fälschl. auch

s**piced **p**ork **a**nd **m**eat/ha**m). Ein Kunde fragt nach einem Gericht ohne Spam, woraufhin der Kellner immer wieder die Karte mit den Spam-Gerichten vorliest, ein Wikinger-Chor fällt mit einem Loblied auf Spam ein, bis das Spam-Gekreische jede Unterhaltung erstickt. Somit lässt sich auch leicht die Herkunft des Wortes für erwünschte Email erklären: Ham (engl. für *Schinken*).

Formen des Spam

Allgemein nennt man das Versenden unverlangter Botschaften an eine Vielzahl von Internet-Nutzern *Spamming*; derjenige, der diese unerwünschten Botschaften verschickt, wird als *Spammer* bezeichnet. Man unterscheidet dabei mehrere Formen des *Spamming*:

Ursprünglich wurden damit unpassende und/oder kommerzielle Beiträge in Newsgroups bezeichnet. Am häufigsten ist damit mittlerweile das Zumüllen eines elektronischen Briefkastens mit **unerwünschten Werbebotschaften** gemeint (UCE).

Daneben gibt es noch das so genannte **Suchmaschinen**- oder **Index-Spamming**, bei dem ein Spammer die Ergebnisse, die eine Internet-Suchmaschine auf eine Stichworteingabe hin ausgibt, mit speziellen Tricks derart manipuliert, dass auf den vordersten Plätzen Webseiten angezeigt werden, die keine für den Surfer relevanten Informationen enthalten.

Spam betrifft auch das Handy: einerseits durch verstärkten Einsatz von Mobile Marketing zur Marktforschung, andererseits durch unerwünschte SMS, die in Japan schon bis zu 90% des Spams ausmachen.

Spammer finden immer neue Techniken, um ihre Werbebotschaften zu verbreiten. So gibt es spezialisierte Programme für fast jeden über das Internet öffentlich zugänglichen Kommunikationskanal: vom Chat wie IRC und ICQ, über den Windows Messenger "Spim" bis hin zu den Logfiles von Webservern, die mit gefälschten Referrer-Daten gefüttert werden.

Spam wird nicht durch den Inhalt definiert, sondern dadurch, dass die Botschaft unverlangt bzw. unpassend ist. Bei den unerwünschten E-Mails unterscheidet man

- UBE (unsolicited bulk e-mail) und
- UCE (unsolicited commercial e-mail).

Kosten

Durch den massenhaften Versand entstehen für Internet-Provider und Unternehmen (durch Bindung der Aufmerksamkeit der Mitarbeiter) beträchtliche Kosten. So soll mehr als die Hälfte der E-Mails, die bei dem amerikanischen E-Mail-Provider *Hotmail* eintreffen, unerwünschte Werbung enthalten. Für den Versand werden Adresslisten genutzt, die oft mehrere Millionen E-Mail-Adressen enthalten.

GESETZE

Nach der EU-Datenschutzrichtlinie, die Ende 2003 in Kraft getreten ist, darf E-Mail-Werbung nur nach vorheriger Zustimmung des Empfängers erlaubt sein.

In Deutschland ist das Versenden von Werbemails ohne Zustimmung des Empfängers wettbewerbswidrig. Das hat der Bundesgerichtshof am 11. März 2004 unter dem Aktenzeichen: I ZR 81/01 entschieden. Allerdings können nur Unternehmen klagen, die im Wettbewerb zu dem Verursacher stehen.

In Österreich war von 1999 bis 2003 für das Versenden von Massen- oder Werbe-E-Mail nach § 101 Telekommunikationsgesetz (TKG) 1997 die vorherige Zustimmung des Empfängers erforderlich *(opt in)*, UCE und UBE somit verboten. Die Nachfolgeregelung, § 107 TKG 2003, erlaubt UCE an Unternehmen oder Behörden, mit Einschränkungen auch an bestehende Privatkunden, wenn diese weitere Nachrichten ablehen können *(opt out)*. Massen- oder Werbe-E-Mail an Privatpersonen bedarf weiterhin der vorherigen Zustimmung des Empfängers *(opt-in)*. Zuwiderhandlungen werden von der Fernmeldebehörde mit bis zu 37.000 Euro bestraft, allerdings ist nur eine Verfolgung österreichischer oder deutscher Täter erfolgversprechend. Unabhängig davon besteht die Möglichkeit einer Klage durch den Empfänger auf Unterlassung oder durch einen Mitbewerber wegen unlauteren Wettbewerbs.

GEGENMASSNAHMEN

Statische Maßnahmen des Filterns basieren immer noch auf der Fähigkeit des Administrators, die vorgegebenen (fixen) Regeln den aktuellen Verhältnissen anzupassen. *Statistisches Filtern*, zuerst vorgeschlagen 1998 am *AAAI-98 Workshop on Learning for Text Categorization* und weiter bekanntgemacht durch einen einflussreichen Artikel von Paul Graham, soll vorhersagen, ob eine E-Mail Spam ist oder nicht. Durch vorheriges Training mit den eigenen E-Mails, die man von Hand in Spam oder Nicht-Spam ("ham") sortiert, lassen sich bereits Aussagen darüber treffen, ob eine E-Mail zum Spam zu rechnen ist oder nicht. Statistische Gegenmaßnahmen basieren auf Wahrscheinlichkeits-Methoden, abgeleitet vom Bayes-Theorem. Bayes'sche Filter sind "lernend" (auch "selbstlernend") und setzen auf Worthäufigkeiten in bereits vom Benutzer erhaltenen und klassifizierten E-Mails. Das Filtern auf statistischen Grundlagen ist im Grunde eine Art Text-Klassifikation. Eine Anzahl von Forschern der angewandten Linguistik, die sich mit dem "Lernen von Maschinen" befassen, haben sich bereits diesem Problem gewidmet. In jüngerer Zeit versuchen Spammer durch das Einfügen zufälliger Zitate aus der Weltliteratur (evtl. in weißer Schrift oder als Meta-Tag unlesbar) die statistischen Maßnahmen auszutricksen.

E-MAILADRESSEN

Eine der effizientesten Maßnahmen der Spamvermeidung für Privatleute besteht darin, die eigene E-Mailadresse nur an engere Bekannte und Freunde weiterzugeben und nicht im Web oder öffentlichen Foren, die von Spammern ausgewertet werden, zu publizieren. Wird für ein öffentliches Forum, zum Beispiel Usenet, eine E-Mailadresse benötigt, lohnt es sich,

Wegwerf-E-Mailadressen mit einem internen Zähler und einer zeitlich beschränkten Gültigkeit anzulegen.

Da die E-Mailadressen aus dem Internet von Robot-Programmen automatisch aus den Newsgroups und Webseiten extrahiert werden, kann die eigene E-Mail-Adresse dort auch so manipuliert werden, dass sie nur von Menschen, aber nicht von Maschinen verstanden werden. Beispielsweise wird statt "Paul@example.org" die Adresse "PaulXYZ@example.org (entferne XYZ)" angegeben. Das Robot-Programm erkennt die Manipulation nicht - die E-Mail-Adresse "Paul@example.org" bleibt Spam-frei.

Allerdings wird z. T. die Ansicht vertreten, die obige Maßnahme bekämpfe nicht die Ursachen, sondern treffe lediglich unbeteiligte Dritte. Das Verwenden ungültiger Adressen empfinden jedoch auch viele als akzeptabel, sofern einige Punkte beachtet werden.

Im Usenet und auf Mailinglisten kann auch im "From"-Header eine nicht gelesene "Müll-Adresse" und "Reply-To" die eigentliche Adresse eingetragen werden. Damit kommen Antworten an der korrekten Adresse an, die Spammer scannen aber normalerweise nur die From-Adressen.

E-Mail-Filter

Manche E-Mail-Provider bieten bereits umfangreiche Dienstleistungen zum Abwehr von Spam, als Beispiel die deutschen Provider GMX und Web.de. Die aktuellen Versionen der E-Mail-Clients Mozilla, Eudora, Opera und Apple Mail haben bereits eingebaute Spamfilter.

Windows-Benutzer können außerdem den E-Mail-/Newsserver "Hamster" einsetzen. Dieser enthält eine mächtige Filtersprache auf Basis von regulären Ausdrücken. Fertige Filterregeln werden auf verschiedenen Seiten einsatzbereit zur Verfügung gestellt, so dass auch Anfänger keine Schwierigkeiten haben, Spam bereits auf dem Server zu löschen, ohne dass die E-Mails komplett geladen werden müssen.

Unter Unixen (z. B. Linux) kann dasselbe durch hintereinander schalten von SpamAssassin (Spamregelwerk) und procmail (Mailfilter) erreicht werden. Bei Spamassassin handelt es sich um ein Programm, welches E-Mails nach bestimmten Mustern durchsucht, die in Werbe E-Mails vorkommen bzw. fehlen. Jedes derartige Muster wird mit einer Zahl (Score) bewertet. Die Bewertungszahlen werden zusammengezählt. Überschreitet die Summe einen bestimmten Wert, wird die Mail als Spam markiert. Diese heuristische Methode kann durch den Bayes-Filter erweitert werden: dieser lernt von dem neu eintreffenden Spam immer weitere "Spam-Schlagwörter" dazu, die beim nächsten Mail mit einem neu angepassten Score in die Bewertung eingehen. Filter wie Bayes-Filter müssen aber zuerst auf einige hundert Emails und Spams trainiert werden, da jeder Benutzer unterschiedliche Emails empfängt. Durch das Training erreicht der Filter aber auch eine höhere Wirksamkeit, die technisch noch nicht einmal voll ausgereizt ist.

Das entscheidende Risiko besteht für den User, dass ihm ein reguläres Mail durch die Lappen geht, also die falsch-positive Fälle.

Massnahmen für Mailserverbetreiber

Kann der einzelne Benutzer nur verhindern, dass er selbst Spam erhält, bietet sich für Administratoren von Mailservern die Möglichkeit, die Verbreitung von Spam einzuschränken. Dies beginnt bei der richtigen Konfiguration des Mailservers, der es nur autorisierten Benutzern gestatten sollte, E-Mails zu verschicken.

Auf der Gegenseite kann der Mailserver den Empfang von E-Mails, die von so genannten Open relays stammen, über die jeder unautorisiert Mails einliefern kann, ablehnen. Mehrere Organisationen, zum Beispiel die Open Relay Database, bieten Listen solcher fehlkonfigurierter Mailserver an, die der Serveradministrator zur Überprüfung nutzen kann.

So genannte Teergruben bieten keinen direkten Schutz vor Spam. Sie bieten eine Gegenmaßnahme gegen den Versandmechanismus der Spammer, indem sie an Hand von Blacklisten den Mail Transfer Agent blockieren.

White/Blacklist Filter bieten einen sicheren Schutz vor Spam, erfordern allerdings die Kooperation legitimer Absender. Das Mailsystem antwortet zunächst allen unbekannten Versendern und fordert diese höflichst auf, sich beim MTA zu registrieren. Durch eine Aktion (z. B. eine Zahl aus einem generierten Bild abschreiben) bestätigt der Sender, dass er ein Mensch ist und ernsthaftes Interesse hat. Wenn er korrekt antwortet, bekommt der Empfänger die bis dahin aufgehobene Mail zugesandt. Der Versender wird daraufhin in die Whitelist aufgenommen. Lehnt der Empfänger den Absender jedoch trotzdem ab, sendet er eine Mail mit dem Subjekt ****SPAM**** an den Absender. Der W/B-Filter fängt diese Mail ab und verschiebt dann die Adresse von der Whitelist auf die Blacklist. Eingehende Mails der Blacklist werden verworfen bzw. automatisch beantwortet. Es gibt noch weitere Registrierungsmöglichkeiten im W/B-Filter- Verfahren, z. B. über einen URL mit ID (z. B. http://www.example.com/mail.php?ID=20032311-021).

Systeme der Art, die die Reaktion des Sendenden erfordert, werden auch als **Challenge-Response-System** bezeichnet, werden jedoch von vielen Anwendern und (vor allem) von Administratoren als kein zweckdienliches System zur Spamvermeidung angesehen. Dies aus den folgenden Gründen:

- Die Absenderadresse einer Spam-Mail wird im günstigsten Fall mit einer ungültigen Adresse, im Normalfall mit der Adresse eines Unbeteiligten versehen. Im Falle einer ungültigen Adresse führt der Versuch der Zustellung der Challenge-Mail zu einem Bounce, damit also zu einer Ressourcenverschwendung. Ist die Adresse gültig, so wird dieser vom Challenge-Response-System "belästigt", womit der Benutzer des Systems technisch selbst zum Spammer wird.
- Versendet der Benutzer eines Challenge-Response-Systems selbst eine Mail an ein Challenge-Response-System (z. B. eine Mailingliste mit Confirmed Opt-in), kommt es zu dem Effekt, dass beide Systeme jeweils auf die Antwort des anderen Systems warten (die Mailliste auf die explizite Bestätigung, dass die Emailadresse in die Liste aufgenommen werden soll, das System des Benutzers, dass sich die Mailliste als "regulärer" Benutzer authentifiziert). Die Aufnahme eines solchen Benutzers erfolgt dann meist

durch manuelles Bearbeiten des Maillistenbetreibers, was für diese einen entsprechenden Mehraufwand bei der Administration zur Folge hat.

- Ein Benutzer eines CR-Systems, der an einer Mailliste teilnimmt, verursacht im Allgemeinen eine Vielzahl von Challenge-Mails, da die Absenderadresse bei Mails an die Mailliste im allgemeinen nicht verändert wird. Dies hat zur Folge, dass sich jeder Maillistenbeteiligte bei jedem einzelnen Benutzer eines solchen Systems authentifizieren muss, damit dieser die jeweilige Mail von der Mailliste erhalten kann. Da dies ab einer gewissen Anzahl von Benutzern von CR-Systemen innerhalb einer Mailliste die Akzeptanzschwelle vieler Benutzer überschreitet, führt dies im allgemeinen dazu, dass sich die Benutzer solcher Systeme früher oder später aus den Diskussionen ausschließen.

Ausblick

Im Kampf um/gegen den Spam wird von beiden Seiten ein immer größer werdender Aufwand getrieben:

- Das Spamaufkommen stieg in den letzten Jahren exponentiell an. Im Jahr 2003 überstieg das Spamaufkommen erstmals die Menge der regulären Mails, so eine Meldung von spamhaus.org Ende des Jahres.
- Aufkommende neue Filter- oder andere Techniken zur Spamvermeidung werden durch entsprechende Gegenmaßnamen umgangen:
    - Die Überprüfung der Gültigkeit von Absenderadressen führte zur Verwendung gültiger Adressen mit dem Effekt, dass Unschuldige mit Tausenden bis zu Millionen von Bounces überschüttet wurden.
    - Die Einführung von Filtern, die Mails auf bestimmte Begriffe überprüften, führte zu Mails, die absichtliche Schreibfehler enthielten (z. B. „V1@gra“ statt „Viagra“) oder durch ungültiges HTML (das von HTML-darstellenden Mailreadern ignoriert wird) den wahren Inhalt verschleierten.
    - Das Sperren bekannter offener Relays und bekannter spamversendender Server führte zur Verbreitung von Trojanischen Pferden, die die Rechner von regulären Benutzern als Spamversender umfunktionierten.
    - Das Einführen von zentralen Listen, die Informationen über offene Relays u. a. verbreiteten und immer öfter von Mailbetreibern genutzt werden, führte zu DOS-Angriffen gegenüber den Betreibern der jeweiligen Liste und deren ISPs
    - Es wird vermutet, dass das 2003 vermehrte Aufkommen von Würmern auf das sich Durchsetzen von statistischen Analysetools (z. B. Bayes-Filtern) zurückzuführen ist.
- Neue Übertragungsmethoden von Mail, die eine Authentifizierung der beteiligten Mailserver erlauben, sollen das bisherige System (SMTP) ablösen. Neben der Ausarbeitung eines neuen Standards von Seiten der IETF, arbeiten große Mailanbieter an eigenen Lösungen. Das Sender Policy Framework ist ein sehr vielversprechendes Kon-

zept, das auf einem zusätzlichen DNS TXT Eintrag basiert. Es werden bereits patches für viele populäre sogenannte MTAs (Mail Transfer Agents) angeboten.

Ein weiterer Ansatz ist die Einführung von E-Mail-Porto. Porto nicht im Sinne von Geld, sondern von Rechenzeit: Wer pro versandter E-Mail 10 Sekunden Rechenzeit zur Verfügung stellen muß, kann nicht unmäßig viele E-Mails verschicken. Allerdings verwenden Spammer dazu ohnehin schon lange nicht mehr ihre eigenen Rechner, weshalb der Erfolg dieser Methode fraglich ist.

Die Erfahrung der letzten Jahre und auch die Tatsache, dass soziale Probleme nicht durch technische Ansätze gelöst werden können, lassen möglicherweise vermuten, dass das System E-Mail in dieser Form in absehbarer Zukunft nicht mehr länger bestehen wird.

Weblinks

- http://www.antispam.de/
- http://www.paulgraham.com/ - Ausführliche Information auf Englisch
- http://www.dr-ackermann.de/spam/faq.htm - FAQ zur Abwehr von Spam
- http://spambayes.sourceforge.net/index.html Spambayes - Englische Antispam-Software die sich in Outlook integrieren lässt
- http://www.spamassassin.org/ - Ein sehr beliebter Bayes-Filter

Wichtige Organisationen

ICANN

Die **Internet Corporation for Assigned Names and Numbers** (ICANN) wird manchmal als eine Art „Weltregierung des Internets“ bezeichnet. Sie verwaltet Namen und Adressen im Internet und koordiniert somit technische Aspekte des Internet.

Die ICANN wurde im Oktober 1998 von einem Zusammenschluss verschiedener Interessenverbände (Wirtschaft, Technik, Wissenschaft und Nutzer) gegründet. ICANN hat die Verantwortung für eine Reihe technischer Vorgaben, die zuvor von der IANA und verschiedenen anderen Gruppen getragen wurden.

Damit das Internet funktioniert, dürfen bestimmte Namen und Adressen weltweit nur einmal vergeben werden. Deshalb koordiniert die ICANN die:

- Internet Domain Namen,
- IP-Adressen,
- Protokoll Parameter und Port-Adressen.

Zudem koordiniert sie den Betrieb der Root-Server des Domain Name Systems, also dem zentralen Adressbuch des Internets.

Das Direktorium der ICANN besteht aus 18 Mitgliedern aus aller Welt. Den Vorsitz hat der Amerikaner Vinton G. Cerf.

Das ICANN ist unter http://www.icann.org zu erreichen.

Internet Engineering Task Force

Die **Internet engineering task force** (IETF) ist neben der Internet Research Task Force (IRTF) eine von zwei Arbeitsgruppen des Internet Architecture Board (IAB). Sie ist eine offene internationale Vereinigung von Netzwerktechnikern, Herstellern und Anwendern die für Vorschläge zur Standardisierung des Internets zuständig ist.

Zurzeit gliedert sich die IETF in 9 Bereiche:

1. Anwendungen (APP)
2. Internet-Dienste (INT)
3. IP:Nächste Generation IPNG
4. Netzwerkmanagement (MNT)
5. Betrieb (OPS)
6. Routing (RTG)

7. Sicherheit
8. Transportdienste (TSV)
9. Benutzerdienste (USV)

Jeder Bereich besitzt zwei Direktoren.

Die Vereinigung wurde 1986 gegründet und kümmert sich im Gegensatz zur IRTF mehr um die kurzfristige Entwicklung des Internets. Es existieren über 80 Arbeitsgruppen mit über 700 Mitgliedern. Eine Gesamtübersicht findet sich im RFC *The TAO of IETF* RFC 3160 (*http://www.ietf.org/rfc/rfc3160.txt*).

Die IETF ist uner http://www.ietf.org erreichbar.

Internet Research Task Force

Die **Internet Research Task Force** (IRTF) ist neben demInternet Engineering Task Force (IETF) die zweite Arbeitsgruppen des Internet Activities Board (IAB). Sie wurde 1998 gegründet um die Forschung und Entwicklung im Bereich der Netzwerke und deren Techniken zu fördern. Sie besteht aus Forschern im Bereich der Netzwerktechnik mit dem Schwerpunkt Internet.

Die Internet Research Steering Group (IRSG) leitet und koordiniert die Forschungsarbeiten. Dabei kommt es mitunter zu Schnittstellen mit den Arbeiten des IETF. Sogar bei den Mitgliedern der Gruppen gibt es Überschneidungen.

Die IRTF besteht aus 12 Forschungsgruppen, die sich unter anderem mit folgenden Themen befassen:

1. End-to-End
2. Information Infrastructure Architecture
3. Privacy and Security
4. Internet Resource Discovery
5. Routing
6. Services Management
7. Reliable Multicast

Nähere Informationen findet man im RF 2014 *IRTF Research Group Guidelines and Procedures* (*http://www.ietf.org/rfc/rfc2014.txt*).

Die IRTF ist unter http://www.irtf.org erreichbar.

RFC

Die **Requests for Comments** (kurz **RFCs**; zu Deutsch etwa 'Bitten um Kommentare') sind eine Reihe von technischen und organisatorischen Dokumenten zum Internet (ursprünglich ARPANET), die am 7. April 1969 begonnen wurde. Bei der ersten Veröffentlichung noch im ursprünglichen Wortsinne zur Diskussion gestellt, behalten RFCs auch dann ihren Namen, wenn sie sich durch allgemeine Akzeptanz und Gebrauch zum Standard entwickelt haben.

Humor in RFCs

Zwischen den offiziellen RFCs, die Quasi-Standards oder „Best Practices“ (derzeit beste Vorgehensweisen) beschreiben, finden sich aber auch immer RFCs, die nicht buchstabengetreu genommen werden sollten, oft aus Anlass des 1. April oder weil gerade Weihnachten war.

- So wird beispielsweise in RFC 2324 ein *Hypertext Coffee Pot Control Protocol (HTCPCP)* definiert, das der Fernsteuerung und -überwachung von Kaffeemaschinen dient.
- Als Parodie auf das Routing-Protokoll MPLS findet sich in RFC 3251 das *Mostly Pointless Lamp Switching*.
- RFC 2795 beschreibt, wie eine unendliche Anzahl von Affen koordiniert werden kann, die die Werke von Shakespeare produzieren soll. Etwas das sogar in ein Programm umgesetzt wurde.
- Aber auch echte Kunstwerke lassen sich ausmachen, so z. B. eine Lobeshymne auf das ARPANET (RFC 527) oder *The 12 Days of Christmas* aus der Sicht eines gestressten Netzwerk-Admins (RFC 1882).
- Am 1. April 2003 ein RFC (RFC 3514) veröffentlicht, das dazu aufruft, bei IP-Paketen, die in irgendeiner Form „Evil“ (böse) sind, ein entsprechendes Bit im Header zu setzen, um diese Pakete an Firewalls leichter ausfiltern zu können.

Nicht immer jedoch bleibt es bei RFCs zum 1. April bei der Theorie. So wurde am 6. März 2001 eine Implementierung des RFC 1149 „*A Standard for the Transmission of IP Datagrams on Avian Carriers*“ (die Übertragung von IP Datagrammen per Brieftaube) vorgestellt. Die durchschnittliche Antwortzeit eines Pings betrug jedoch 45 min, so dass nicht mit einer regelmäßigen Nutzung im Echteinsatz zu rechnen sein wird.

Einige RFCs sind zugleich *For Your Information (FYI)* mit eigener Zählung. So ist der RFC 1462 identisch mit dem FYI 20 *FYI on What is the Internet?*. Ebenso sind einige RFCs auch *Best Common Practice (BCP)* oder *RARE Technical Report (RTR)*.

Die RFCs sind unter http://www.rfc-editor.org/ einsehbar.

World Wide Web Consortium

Das **World Wide Web Consortium**, oder auch **W3C**, ist das Gremium zur Standardisierung das Internet betreffender Techniken. Gründer und Vorsitzender des W3C ist Tim Berners-Lee, der auch als der Erfinder des World Wide Web bekannt ist.

Beispiele für bisher vom W3C verabschiedete Standards sind HTML, XML, CSS und WAI. Das W3C und seine Mitglieder beschäftigen sich auch mit der Weiterentwicklung von Standards oder mit der Entwicklung neuer Standards.

Bei der Entwicklung neuer Standards hat sich das W3C verpflichtet, nur noch Technologien zu verwenden, die frei von Patentgebühren sind.

Das W3C ist unter http://www.w3.org/ zu erreichen.

Internetarchiv

Das **Internetarchiv** in San Francisco wurde 1996 von Brewster Kahle gegründet und speichert Momentaufnahmen von Webseiten, Usenetbeiträge, Filme, Tonaufnahmen (hauptsächlich von Live-Konzerten), Bücher und Software. Eine Sicherungskopie der Daten von San Francisco befindet sich in der Bibliotheca Alexandrina.

Zum Archiv gehört auch die **Wayback Machine** (Zeitmaschine), mit der man die verschiedenen gespeicherten Versionen abrufen kann. Der Gesamtumfang beträgt heute (2004) etwa 30 Milliarden Seiten. Die Seiten werden erst 6 Monate nach dem indexieren öffentlich verfügbar gemacht.

Das Internetarchiv ist unter http://web.archive.org zu erreichen.

DENIC

Die **DENIC eG** ist die zentrale Registrierungsstelle für Domains unterhalb der Toplevel-Domain DE. 'DE' ist der Ländercode für Deutschland nach ISO 3166. Die DENIC eG ist eine eingetragene Genossenschaft mit 80 Mitarbeitern und wurde am 17. Dezember 1996 gegründet. Die Mitglieder der DENIC eG sind Internetdienstanbieter, die ihren Kunden lokale Zugänge zum Internet zur Verfügung stellen.

Zu den Aufgaben der DENIC eG gehören:

- Betrieb des ersten Rootservers in Deutschland (K-Rootserver Kopie)
- Betrieb des Primary-Nameservers für die Toplevel-Domain DE
- Bundesweit zentrale Registrierung und Verwaltung von Domains unterhalb der Toplevel-Domain DE

- Administration des Internet in Zusammenarbeit mit internationalen Gremien (CENTR, ICANN, CORE)
- Bereitstellung verschiedener Datenbankdienste
- Bereitstellung verschiedener Informationen, insbesondere zu rechtlichen Fragen bei der Domainregistrierung und -verwaltung

Um die Bemühungen zu einer stärkeren weltweiten Verteilung der DNS-Rootserver zu unterstützen, beteiligen sich die DENIC eG als zentrale Registrierungsstelle für .de-Domains und der Internet-Verband eco als Betreiber des Netzknotens DE-CIX an der Einrichtung und dem Betrieb eines Nameservers in Frankfurt am Main, eine exakte Kopie des K-Rootservers in London. Der Server ging Anfang 2004 ans Netz, und hat dadurch die Zugriffsgeschwindigkeit von DNS-Abfragen beschleunigt. Sowohl diese Kopie, als auch die Rootserver in London und Amsterdam selbst werden von RIPE betreut. Die DENIC sorgt für den netztechnischen Zugang und sponsert die benötigte Hardware, während eco die Örtlichkeiten zur Verfügung stellt und für einen Vor-Ort-Service rund um die Uhr Sorge trägt.

Am 1. März 2004 führte die DENIC Umlautdomains im deutschen Sprachraum ein.

Die Denic ist unter http://www.denic.de/ zu erreichen.

Chaos Computer Club

Der **Chaos Computer Club** (CCC) ist ein deutscher Verein von und für Hacker. Wichtigste Ziele sind *Informationsfreiheit* und ein *Menschenrecht auf Kommunikation.* Jedem, der sich mit diesen Zielen identifizieren kann, steht die Mitgliedschaft offen. Dabei herrscht eine gewisse Dissonanz bezüglich der gewünschten Organisationsform. Einerseits sieht man sich als „galaktische Gemeinschaft“, die nicht auf plumpe Verwaltungsakte angewiesen ist, andererseits gibt es einen eingetragenen Verein, der ca. 1500 Mitglieder zählt.

Struktur und Veranstaltungen

Der CCC e.V. ist dezentral in einzelnen regionalen Gruppen organisiert. Kleinere Gruppen heißen Chaostreffs, während aktivere und größere sich ERFA-Kreise (Erfahrungsaustauschkreise) nennen. Der erste virtuelle ERFA-Kreis sind die Haecksen, die weiblichen Mitglieder des CCC.

Mitglieder und Interessierte treffen sich seit 1984 einmal jährlich zum *Chaos Communication Congress*. Außerdem fand im Sommer 1999 und 2003 das *Chaos Communication Camp* auf dem *Paulshof* nahe der Kleinstadt Altlandsberg auf dem Land statt. Im Gegensatz zum Congress ist es eine internationale Veranstaltung und die Konferenzsprache deshalb Englisch. Neben den vielen Vorträgen über technische und gesellschaftspolitische Themen gibt es auch Workshops über z. B. das Lockpicking. Zu Ostern findet regelmäßig der workshoporientierte Easterhegg statt.

Publikationen

Der CCC gibt etwa vier Mal jährlich die Zeitschrift Die Datenschleuder heraus. Von der Hackerbibel, einem umfangreichem Kompendium und Sammelsurium von zahlreichen Dokumenten der Hackerszene, erschienen in den 1980er Jahren zwei Ausgaben.

Des weiteren wird auf Radio Fritz einmal im Monat die Sendung Chaosradio ausgestrahlt. Eine weitere Radiosendung des CCC ist der in Darmstadt beheimatete C-Radar. Wichtigste Kommunikationsmittel sind Mailinglisten und Webseiten.

Geschichte

Gegründet wurde der CCC 1981 in Berlin am Tisch der Kommune 1 in den Redaktionsräumen der taz. Der eingetragene Verein wurde ein paar Jahre später in Hamburg gegründet. In der Folge entwickelte der Club vor allem in Hamburg, bevor er später durch weitere sogenannte ERFA-Kreise und Chaostreffs sich auch auf andere Städte ausdehnte.

Öffentliche Bekanntheit erlangte er 1984 mit dem BTX-Hack, der später im Fernsehen nachgestellt wurde, und dem NASA-Hack drei Jahre später. Wenig später geriet der Verein ins Zwielicht, als Hacker aus dem Umfeld des CCC, zu nennen ist vor allem Karl Koch, Informationen an den KGB verkauften (KGB-Hack).

Ein weiteres düsteres Kapitel ist der Tod des Hackers Tron, der 1998 erhängt aufgefunden wurde. Manche Mitglieder des CCC vertreten vehement eine Mordtheorie. Die Umstände des Falls konnten bislang nicht aufgeklärt werden. Im Jahr 2001 starb Wau Holland, Gründer und Vaterfigur des Chaos Computer Club.

Ebenfalls im Jahr 2001 feierte der Club sein 20-jähriges Bestehen mit einer interaktiven Lichtinstallation namens Blinkenlights am Haus des Lehrers am Alexanderplatz (Berlin).

Bekannte CCC-Mitglieder sind unter anderem der Gründer Wau Holland, Steffen Wernéry, Andy Müller-Maguhn, der von 2000 bis 2002 einen Sitz im Direktorat der ICANN hatte, und Peter Glaser.

Häufig arbeitet der CCC auch mit anderen Organisationen, die sich gegen Zensur, für Informationsfreiheit oder den Datenschutz einsetzen, zusammen. Insbesondere sind hier der FITUG und der FoeBuD zu nennen.

Literatur

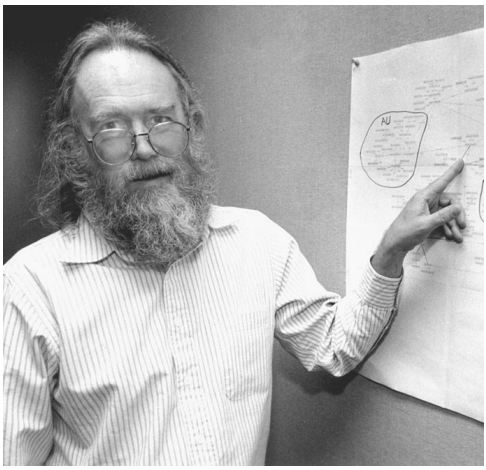
- Daniel Kulla: *Der Phrasenprüfer. Szenen aus dem Leben von Wau Holland, Mitbegründer des Chaos Computer Clubs*, Löhrbach 2003, ISBN 3922708250

Der CCC ist unter http://www.ccc.de zu erreichen.

Perönlichkeiten

Jonathan Postel

Jonathan *Jon* Bruce Postel (* 6. August 1943; † 16. Oktober 1998 in Santa Monica) war Internet-Pionier.Während seines Studiums an der University of California Los Angeles (wo er 1974 promovierte) war er als Assistent am ARPAnet beteiligt. Er war maßgeblich beteiligt an der Entwicklung der Addressierung der Computer im Internet.



Er war Gründer der Internet Society und Herausgeber des Request for Comments das ihm nach seinem Tode mit dem RFC 2468 einen virtuellen Gedenkstein setzte.

Der Internet Assigned Numbers Authority stand er als Direktor vor. Er war Gründungsmitglied des Internet Architecture Board.

Mit seinem buschigen Bart, langen Haaren und freundlichem Wesen hatte er eine sehr bedeutende Stellung inne, die der Öffentlichkeit aber eher unbekannt war.

Für seine Arbeit an der Entwicklung des Internets wurde er kurz vor seinem Tod von der International Telecommunication Union ausgezeichnet.

Robert E. Kahn

Robert E. Kahn ist Informatiker und zusammen mit Vint G. Cerf einer der Entwickler des Transmission Control Protocol (TCP), das im modernen Internet zur Datenübertragung dient.

Er erreichte 1960 am City College of New York einen Bachelor-Abschluss und an der Universität Princeton einen M.A (1962) und Ph.D. (1964). Er arbeitete anschließend in den Bell Laboratories und am MIT als Assistenzprofessor für Elektrotechnik. Er ließ sich vom MIT beurlauben, um bei Bolt Beranek and Newman zu arbeiten, wo er für das Systemdesign des ARPANETs zuständig war.

1972 wechselte er zur Defense Advanced Research Projects Agency (DARPA), und im Oktober desselben Jahres demonstrierte er auf der *International Computer Communication Conference* das ARPANET mit 40 verbundenen Computern, die erste Präsentation vor der Öffentlichkeit. Nach diesem Erfolg wurde er Direktor des Information Processing Techniques Office (IPTO) des DARPA. In dieser Position startete er ein eine Milliarde US-Dollar

teures Programm der USA, das bis dahin größte staatliche Projekt in der Computerforschung und -Entwicklung.

Er hatte die grundlegenden Ideen für das **Transmission Control Protocol**, als er an einer Datenübertragung per Satellit arbeitete. Während er an diesem Projekt arbeitete, erstellte er die Basis für eine offenen Netzwerkarchitektur, die es unterschiedlichen Netzwerken erlaubte, miteinander zu kommunizieren. Die Art der Hardware oder Software in den einzelnen Netzen spielte dabei keine Rolle. Um diese Anforderungen zu erfüllen, entwickelte er TCP nach folgenden Eigenschaften:

- Kleine Sub-Netzwerke des Hauptnetzes können miteinander durch eine Art Gerät kommunizieren. Daraus entwickelen sich später Gateways.
- Kein Punkt des Netzwerkes darf bei Fehlern für den Zusammenbruch des ganzen Netzwerks verantwortlich sein. Redundanz wird also zur Regel.
- Wenn Informationen bei der Übertragung von einem Computer zum anderen verloren gehen, werden sie nochmals gesendet.
- Jeder Computer kann mit dem Netzwerk verbunden werden ohne das Netzwerk intern zu verändern.

Ab Frühling 1973 unterstützte ihn Vinton G. Cerf bei dem Projekt, und zusammen schlossen sie das Projekt TCP erfolgreich ab. Dabei diente ihnen das **Networking Control Protocol** als Grundlage. Später kam zu TCP noch das **Internet Protocol** (IP) hinzu und beide, zusammengefasst unter der Bezeichnung TCP/IP, sind die Grundlage des modernen Internets.

Nach 13 Jahren verließ er DARPA und gründete die **Corporation for National Research Initiatives** (CNRI), deren Vorsitzender er ist. Die CNRI ist eine gemeinnützige Organisation, um Entwicklung und Forschung der Informations-Infrasturktur voranzutreiben.

Vinton Gray Cerf

Vinton *Vint* Gray Cerf (* 23. Juni 1943 in New Haven) ist Mathematiker und Computerwissenschaftler.

Vint war schon in seiner Kindheit ein guter Mathematiker und auch an der Chemie interessiert.

Cerf wird oft als einer der Väter des Internets bezeichnet. Während seiner Arbeit bei der US-amerikanischen DARPA zwischen 1976 und 1982 spielte er eine Schlüsselrolle in der Entwicklung des Internets und der im Internet verwendeten Verbindungsprotokolle. So hat er zum Beispiel zusammen mit Robert E. Kahn das Transmission Control Protocol.

Als Vize-Präsident von **MCI Digital Information Services** zwischen 1982 und 1986 leitete er die Entwicklung von MCI Mail, der ersten kommerziellen Email-Service im Internet.

Im Dezember 1997 wurde er zusammen mit Robert E. Kahn von US-Präsident Bill Clinton für seine Verdienste mit der **U.S. National Medal of Technology** ausgezeichnet.

Er ist Autor mehrerer RFCs und Gründer der ISOC.

Eric Allman

Eric Allman (* 1959 in El Cerrito, Kalifornien) ist ein Computer-Programmierer. Der offen schwule Programmierer gilt als Vater der modernen Internet-E-Mail.

Während seiner Anstellung an der University of California Berkeley in den 1970ern und 1980ern entwickelte er die Software Delivermail und die auch heute noch sehr wichtige Nachfolgersoftware Sendmail.

Schon mit acht Jahren interessierte sich Allman für Computer und sah darin seine Zukunft. Als Teenager hackte sich in den Zentralrechner seiner High School und nutzte er auch das Rechenzentrum der **Universität Berkeley** für sich.

1973 begann er als Freshman sein Studium an der Universität, im gleichen Jahr in dem Unix von AT&T an die Universität kam. Da der Quellcode von Unix mehr oder weniger frei war, entstanden auch von Anfang an Erweiterungen von Unix und neue Software. Dazu gehörte auch **Delivermail** von Allmman das die Kommunikation per mail verbesserte und den netzübergreifenden EMail-Verkehr im ARPANET ermöglichte. Die Subnetze des ARRPANETs schrieben teils sehr unterschiedliche Formate für EMail vor, die Delivermail aufgrund seiner großen Toleranz gegenüber fehlerhaften EMails beherrschte.

1981 wurde aus Delivermail der bekanntere Name Sendmail und im Jahr darauf wurde es ein wichtiger Teil der **Berkeley Software Distribution**. Auch heute noch wird Sendmail, trotz zahlreicher Kritik aufgrund seiner Fehlertoleranz und damit auch Fehlerhaftigkeit und der kryptischen Konfiguration, immer noch auf vielen Unix und Linux Systemen verwendet.

Nachdem er Berkeley verlassen hatte gründete Allman 1998 die Firma **Sendmail Inc**. in Emeryville, Kalifornien um Sendmail weiter zu entwickeln und neben kommerziellem Support auch kommerzielle Erweiterungen wie eine GUI für die Konfiguration zu bieten.

Zitat

„Es ist eine Art perverse Befriedigung zu sehen dass es im Grunde unmöglich ist Hassmails durch das Internet zu schicken ohne dass diese von einem schwulen Programm berührt wird. Das ist schon lustig.“

Orginal: „*There is some sort of perverse pleasure in knowing that it's basically impossible to send a piece of hate mail through the Internet without its being touched by a gay program. That's kind of funny.“*

Seine persönliche Website findet man unter http://www.sendmail.org/~eric/

TIM BERNERS-LEE

Dr. **Timothy J. Berners-Lee** (* 8. Juni 1955 in London) gilt als Erfinder des World Wide Webs.

Berners-Lee hat an der Oxford University in England promoviert. Derzeit ist er Inhaber des 3Com Founders-Lehrstuhls am *Laboratory for Computer Science* des *Massachusetts Institute of Technology* (MIT). Zudem steht er dem World Wide Web Consortium vor, einem offenen Forum für Unternehmen und Organisationen, das die weitere Entwicklung des WWW begleitet. Für seinen Beitrag zur Wissenschaft, den er durch die Erfindung des World Wide Web leistete, wurde er 2004 zum *Knight Commander, Order of the British Empire (KBE)*, also zum Ritter ernannt.

1989 schlug Berners-Lee seinem Arbeitgeber CERN (Europäisches Labor für Teilchenphysik) ein Projekt vor, das auf dem Prinzip des Hypertexts beruhte und den weltweiten Austausch sowie die Aktualisierung von Informationen zwischen Wissenschaftlern vereinfachen sollte. Er verwirklichte dieses Projekt, das den Ursprung des World Wide Webs darstellt.

Die erste Webseite, die Berners-Lee erstellte (und damit die erste überhaupt), war http://info.cern.ch (die Seite existiert nicht mehr, es gibt aber eine Kopie (*http://www.w3.org/History/19921103-hypertext/hypertext/WWW/TheProject.html*)). Sie erläuterte unter anderem,

- was das Internet sein sollte,
- wie man an einen Webbrowser kommt,
- wie man einen Webserver aufsetzt.

Ursprünglich war dies auch die erste einfache Suchmaschine, denn Berners-Lee betreute noch andere Webseiten außer seiner eigenen.

Die Grundideen des World Wide Webs sind vergleichsweise einfach zu begreifen. Berners-Lee sah und verknüpfte sie jedoch in einer Weise, deren Möglichkeiten bis heute noch nicht vollständig ausgeschöpft sind.

Das wohl wichtigste war aber, dass er seine Ideen und technischen Umsetzungen nicht patentierte sondern frei weitergab. Auch auf die Maxime des World Wide Web Consortiums nur patentfreie Standards zu verabschieden hatte er starken Einfluss.

In seinem Buch *Weaving the Web* (deutsch: *Der Web-Report*, 1999) wird z. B. folgendes betont:

- Das Web editieren zu können ist genauso wichtig, wie durch das Web zu browsen.
- Computer können genutzt werden, um im Hintergrund Aufgaben zu erledigen, damit wir in **Gruppen** besser **zusammenarbeiten** können (hierfür ist ein Wiki ein gutes Beispiel).

- Jeder Bereich des Internets sollte eher eine **Netzstruktur** denn eine **Baumstruktur** haben. Erwähnenswerte Ausnahmen sind das Domain Name System und die Regeln für die Vergabe von Domainnamen durch die ICANN;
- Informatiker tragen nicht nur eine technische, sondern auch eine moralische Verantwortung.

Literatur

- Tim Berners-Lee mit Mark Fischetti: Der Web-Report. Der Schöpfer des World Wide Webs über das grenzenlose Potential des Internets. Aus dem Amerikanischen von Beate Majetschak. Econ 1999, ISBN 3-430-11468-3

Al Gore und das Internet

Albert (*Al*) Gore (* 31. März 1948) ist ein US-amerikanischer Politiker.

Gore ist Mitglied der Demokraten. Von 1984 bis 1993 war er Senator des Bundesstaates Tennessee, vom 20.1.1993 bis zum 20.1.2001 war Gore Vizepräsident unter Präsident Bill Clinton.

Am 9. März 1999 wurde **Al Gore** von **Wolf Blitzer** (CNN) interviewt. Während dieses Interviews sagte Gore: „*During my service in the United States congress I took the initiative in creating the Internet.*“ (dt: „Während meiner Amtszeit im US-Kongress ergriff ich Initiative in der Schaffung des Internets“). Diese Aussage sorgte erst für keine Überraschungen, und keiner der Journalisten, die über das Interview berichteten, beachtete die Aussage.

Erst zwei Tage später begann die **Republikanische Partei** Pressemeldungen herauszugeben, in denen sie Gore für seine Behauptung, das „Internet erfunden zu haben“, anprangerte. Nicht nur in den konservativen Medien, sondern auch bei den Internet-Geeks wurde „***Al Gore said he invented the Internet!*“** schnell ein Renner.

Die Pressemitteilungen enthielten ganz korrekt den Hinweis, dass das ARPANET, der Vorgänger des Internets, schon 1971 bestand, lange bevor Gore für den Kongress kandidierte.

Jedoch war das ARPANET ein relativ kleines privates Projekt von Universitäten, während das Internet ein privatwirtschaftliches Projekt von gigantischem Ausmaß ist. Gores Aussage bezog sich auf ein Gesetz von 1990, das die Entwicklung eines „**Informations System Highway“** für den wissenschaftlichen und erzieherischen Sektor fördern sollte. Das Gesetz hatte damit einen sehr großen Einfluss auf das Wachstum des damals noch kleinen Internet.

Am 28. September 2000 wurde dann von **Vint G. Cerf** und **Robert E. Kahn**, beide wichtige Persönlichkeiten in der Entwicklung des Internets, eine gemeinsame Email verbreitet. Die beiden nahmen Al Gore in Schutz und würdigten seine politische Unterstützung für die Entwicklung des Internets. Sie würdigten auch seine Arbeit als Vizepräsident der USA, in welcher er den staatlichen Einfluss auf das Internet verringerte und die Nutzung des Internets in staatlichen Institutionen und Schulen förderte.

Leider wird Al Gores Aussage ihm immer noch zu negativ angerechnet, aber er nimmt es mit Humor und sagte in der TV-Show von David Letterman „*I gave you the Internet, and I can take it away!*“.

Mensch und Internet

Netzkultur

Netzkultur oder auch **Internetkultur** ist die Kultur des Internets. Für viele Menschen ist das Internet aus ihrem Alltag nicht mehr wegzudenken. Es verändert wie jedes neue Medium auch die Gesellschaft.

Bei der Internetkultur handelt es sich um eine weltweite Subkultur im soziologischen Sinne. Ihre Geschichte spiegelt sich in der Internetfolklore wieder und es existieren mit der Netiquette klare Verhaltensregeln. Daneben sind auch eine ganze Reihe von Insider-Witzen und Running Gags in Umlauf. Das Zusammengehörigkeitsgefühl der Angehörigen dieser Kultur drückt sich unter anderem durch eine eigene Sprache, dem so genannten Netzjargon aus.

Insider der Netzkultur bezeichnen sich selbst als Regulars, im Unterschied zum Newbie (Anfänger) und erst recht von „normalen Menschen“. Neben der Netzwelt gibt es auch noch das reale Leben, Real Life genannt, das sich außerhalb des Netzes und abseits des Computers abspielt: „Draußen, das ist da, wo die Pizza herkommt“ ;)

Literatur

- Lovink, Geert: Dark Fiber - Auf den Spuren einer kritischen Internetkultur, Bundeszentrale für politische Bildung, ISBN 3810041459
- http://www.heise.de/tp – Telepolis Magazin der Netzkultur

Internetsucht

Unter **Internetsucht** versteht man den zwanghaften Drang, jeden Tag oft stundenlang im Internet herumzusurfen. Oft werden bestimmte Foren oder Chatrooms aufgesucht, in denen man sich regelmäßig mit Gleichgesinnten trifft. Aber auch Sex-Seiten und Spielangebote werden fleißig genutzt. Das Internet spielt im Leben der Betroffenen die Hauptrolle. Andere Verpflichtungen, die das Leben normalerweise mit sich bringt, werden meist vernachlässigt. Im Extremfall wird die virtuelle Welt zu einem Ersatz für die sonst üblichen realen sozialen Kontakte. Normale freundschaftliche Kontakte werden nicht mehr getätigt. Es kommt zu einem Kontrollverlust, der den User zwingt, immer länger im Netz zu verweilen. Hinterher haben viele Internetsüchtige oft Schuldgefühle, können sich von ihrer Sucht aber nicht befreien.

Nach außen verheimlichen Internetsüchtige oft ihre Sucht oder wollen sie nicht wahrhaben. Sie bagatellisieren zum Beispiel ihr Verhalten. Ist der PC einmal defekt, kommt es zu Entzugserscheinungen, schlechter Laune, Nervosität, Reizbarkeit, Schlafstörungen und

Schweißausbrüchen. Es kann auch dazu kommen, dass der Süchtige kaum noch etwas isst, wenn er vor dem Monitor sitzt.

Als besonders gefährdet gelten depressive und einzelgängerisch veranlagte Menschen. Besonders verbreitet soll die Internetsucht bei männlichen Surfern unter 18 sein, weil sie sich dem Druck des Alltags nicht gewachsen fühlen und zum Ausgleich in die virtuelle Welt flüchten. Schüler vernachlässigen ihre Hausaufgaben. Erwachsene ziehen sich immer mehr von der Außenwelt zurück. Viele surfen nachts stundenlang herum und kommen übermüdet zur Arbeit.

Historisches

Der Begriff **Internetsucht** (Onlinesucht) wurde von dem New Yorker Psychiater Ivan Goldberg, selbst ein intensiver Internet-Nutzer - zunächst eher scherzhaft - geprägt. Das Thema „Internet Addiction“ (Netaddiction) entwickelte sich schnell zum Gesprächsthema nicht nur in der Internetgemeinde. Die New York Times hatte im Februar 1995 einen Artikel zum Thema „Internetsucht“ geschrieben. Seither häufen sich die Berichte und auch Untersuchungen.

Epidemiologie

Seit dem Jahr 2000 hat sich nach den Untersuchungen des Berliner Psychiaters Werner Platz die Zahl der Internetsüchtigen in Deutschland vervierfacht. In Berlin gibt es ca. 10.000 Internetsüchtige, was aber nur die Spitze des Eisberges sei.

In den USA wird die Zahl der an **Internetsucht**-Erkrankten auf ca. 200 000 geschätzt. Diese Schätzung geht aus einer Studie der amerikanischen Psychologie-Professorin Kimberly S. Young hervor (*http://www.netaddiction.com/*). Young schätzt die Internet-Sucht, die sie „pathological internet use“ (PIU) nennt, weltweit auf etwa 7% der www-Surfer. Eine österreichischen Studie (Zimmerl und Panosch - *http://gin.uibk.ac.at/thema/internetsucht/internetsucht-zus.html*) ergab *dass 12.7% der untersuchten Probanden ein suchtartiges Verhalten aufweisen, welches man als „Pathologischen Internet - Gebrauch (PIG)“ bezeichnen könnte.* Gemäß einer wissenschaftliche Studie der Humboldt Universität Berlin, in der über 7000 Internet-User über ihre Gewohnheiten befragt wurden, verbringen Internetsüchtige (3% der Befragten) durchschnittlich 35, Gefährdete (7%) 29 Stunden Stunden pro Woche im Netz. Der durchschnittliche Normal-User (90%) bringt es auf 17,5 Stunden. Jugendliche und junge Erwachsene sind besonders betroffen. Niedriger sozialer Status, Arbeitslosigkeit und fehlende Partnerschaft sind Risikofaktoren.

Symptomatik

- Häufiges unüberwindliches Verlangen, ins Internet einzuloggen
- Kontrollverluste (d. h. längeres Verweilen „online“ als intendiert) verbunden mit diesbezüglichen Schuldgefühlen

- sozial störende Auffälligkeit im engsten Kreis der Bezugspersonen (Freunde, Partner, Familie), häufige Rügen durch unmittelbare Bezugspersonen
- nachlassende Arbeitsleistung
- Verheimlichung/ Bagatellisierung der Online-Aktivitäten vor der Umwelt
- Psychische Irritabilität bei Verhinderung am Internet-Gebrauch (kann sich auswirken in Form von Nervosität, Reizbarkeit und Depression)
- Mehrfach fehlgeschlagene Versuche der Einschränkung

Stadien

- **Gefährdungsstadium**: Dieses ist gegeben, wenn 3 der oben beschriebenen Kriterien über 6 Monate vorhanden sind.
- **kritisches Stadium**: Hier müssen es bereits 4 Kriterien über einen Zeitraum von 4 Monaten sein.
- **chronisches Stadium**: Dies ist gegeben, wenn das kritische Stadium mit den 4 Kriterien überstiegen wird und bereits Folgeschäden auftreten, z. B. Jobverlust, Trennung von Partnern, Abkapselung von der Familie, Verschuldung oder physische Schäden, z. B. der Augen oder der Wirbelsäule

Emoticon

Ein **Emoticon** ist eine Zeichenfolge, die einen Smiley nachbildet und dazu verwendet wird, in der schriftlichen elektronischen Kommunikation Stimmungs- und Gefühlszustände auszudrücken.

Am 19. September 1982 schlug der Student **Scott E. Fahlman** in einem Bulletin Board (elektronisches Diskussionsforum) der *Carnegie Mellon University* vor, aus ASCII-Zeichen das inzwischen weltberühmt gewordene Signet nachzubilden.

Verwendet werden Emoticons etwa in Chaträumen, im Usenet und im E-Mail-Verkehr.

Die bekanntesten sind wohl

```
:-)  lachendes Gesicht
:-(  trauriges Gesicht
;-)  zwinkern
:p   Zunge rausstrecken
```

Es gibt eine Unzahl von Emoticons. Auch verwenden nicht alle das selbe Emoticon für ein und dieselbe Sache. Ein lachendes Gesicht kann zum Beispiel sein:

```
:)    :o)   :-)   =o)
```

In jüngerer Zeit werden von einer zunehmenden Anzahl von Programmen Emoticons in eine Grafik umgewandelt, etwa von Instant-Messenger-Software oder E-Mail-Clients wie Mozilla Mail.

Weitere Beispiele

(-: Gute Laune (Linkshänder)	:-/ das finde ich nicht lustig
:)) über beide Ohren glücklich	:-> sarkastisches Grinsen
:-o Oh, Oh	>-) freches Grinsen
%-) verwirrt	:'-) Freudentränen
:-& ich bin sprachlos	:*) herumblödeln oder betrunken
:-O überrascht, schockiert	:-x Kuss
:-(traurig, unzufrieden	:-X dicker Kuss
:'-(weinen	:-D laut lachend
>-(wütend	8-) Brillenträger
:-# meine Lippen sind versiegelt	:-Q Raucher
:-| ich sage diesmal nichts	@-;-'- Rose
%-} Blödsinn	B-) Brillenträger

Während bei den traditionellen Emoticons der Betrachter den Kopf nach links neigen muss (oder bei Linkshändersmileys **(-:** nach rechts), kommt eine alternative Form von Emoticons aus Japan, bei der dies nicht nötig ist. Seit dem Jahr 2000 ist diese Form von Emoticons vermehrt im IRC anzufinden und erfreut sich insbesondere bei jungen Menschen besonderer Beliebtheit.

(^_^) männlich	(;_;) weinen
(^.^) weiblich (verbergen der Zähne)	(?_?) erstaunen
\(^_^)/ Hurra! (mit Armen; kleinere Version: \o/)	(>_<) Augen zukneifendes Gesicht - Autsch!
(^_~) Auge zukneifen / zwinkern	(-_-)zzz schlafen
(^_-) Auge zukneifen / zwinkern	(-_-) genervt
(*_*) Angst haben	(`_') böses Gesicht

Oftmals werden aus Gründen der Zeitersparnis die Gesichtsumrandungen weggelassen: ^.^ statt (^.^) Das kürzeste Emoticon dieser Art: ^^

Englische Sprache im Internet

Die Verwendung der **Englischen Sprache im Internet** hat in den letzten Jahren einige Wandlungen erlebt.

Vor allem junge Menschen verwenden für häufig gebrauchte Wörter eine andere Rechtschreibung und wollen dadurch offenbar besonders cool wirken.
Beipielsweise wird das Plural-"*s*" durch ein „*z*" ersetzt wie in Warez, Toolz etc.
Auch die **Ziffern 0-9** erhielten eine größere Bedeutung und manche Buchstaben werden durch die Ziffer, die ihnen am meisten ähnelt ersetzt (sogenannte Leetspeak). So wird aus „*lame*" (langsam/langweilig) eine „*14m3*" (meist 1=L/I; 3=E; 4=A; 5=S; 7=T; 8=B; 9=G).

Zeit ist ein wichtiger Faktor zum Beispiel in Chaträumen, wo ein langsamer Schreiber zur Verwendung von **Abkürzungen** greift. Hier werden Abkürzungen auch verwendet, um seine Gefühle wie Freude schnell auszudrücken.
Hier wird aus einem *As far as I know* plötzlich *AFAIK*.

Häufige Abkürzungen: ebee

Abkürzung	English	Deutsch
ACK	Acknowledge	Zustimmung
AFAIK	As far as I know	Soweit ich weiß
AFAIR	As far as I remember	Soweit ich mich erinnere
BRB	Be right back	Bin gleich zurück
BTW	By the way	Nur so nebenbei
FYI	For your information	zu deiner Information
IMO	In my opinion	Meiner Meinung nach
IMHO	In my humble opinion	Meiner bescheidenen Meinung nach
LOL	Laughing Out Loud	lautes Lachen
ROTFL, ROFL	Rolling on the floor laughing	„sich vor Lachen auf dem Boden wälzen“

NETIQUETTE

Die **Netiquette** (aus engl. *net*, Netz, und *etiquette*, Etikette) ist ein wichtiger Bestandteil der Netzkultur. Sie beschrieb ursprünglich die grundlegenden Verhaltensrichtlinien im Usenet. Der Begriff wird aber mittlerweile allgemein für Verhaltensregeln in Datennetzen wie dem Internet verwendet.

Die erste und grundlegende Regel der Usenet-Netiquette ist: "Vergessen Sie niemals, dass auf der anderen Seite ein Mensch sitzt!" (Wenn man auch nicht sicher sein kann, ob es vielleicht ein "seelenloses" Programm ist, mit dem man interagiert).

Einzelne Bestimmungen der Netiquette, unter anderem die Forderung nach einem Realname, nach der es im deutschsprachigen Usenet als unhöflich gilt, unter einem falschen Namen bzw. Pseudonym zu posten, geraten zuweilen unter Beschuss, wenn aufgrund des Themas oder der Art der Diskussion Anonymität wünschenswert oder notwendig ist.

Die RFC 1855, die *Netiquette Guidelines*, spricht sich deutlich gegen Fullquotes und TOFU aus. Trotzdem gibt es bei manchen Diskussionen den Bedarf von intensivem Zitieren, z.B. wenn der Empfängerkreis vergrößert werden soll, ohne dass alle auf den bisherigen Verlauf zugreifen können.

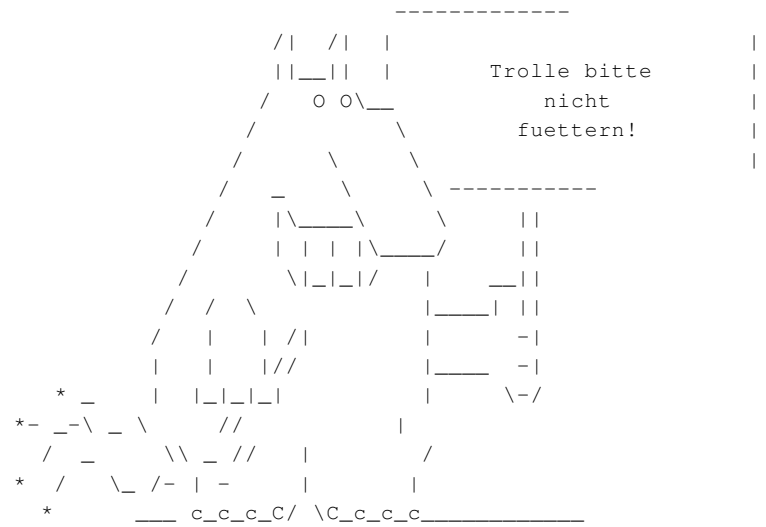
TROLL

Im Internet werden Menschen als **Troll** bezeichnet, die Beiträge verschicken, mit denen sie erkennbar nur provozieren wollen, ohne einen wirklichen Beitrag zur Diskussion leisten zu wollen.

Trolle bewegen sich vor allem in Diskussionsforen bzw. Newsgroups, aber auch in Wikis und auf Mailinglisten. Ihr Ziel ist es, Diskussionen um ihrer selbst willen auszulösen oder zu betreiben ohne wirklich am Thema interessiert zu sein, wütende Antworten auszulösen, Menschen mit anderer Meinung zu diskreditieren oder eine Diskussion zu sabotieren, indem eine unangenehme Atmosphäre geschaffen wird.

Häufig werden jedoch Diskussionsteilnehmer, die lediglich eine kontroverse Meinung äußern, als Troll diskreditiert, um so einer sachlichen Diskussion aus dem Wege zu gehen.

Obwohl Trolle in der Mythologie nicht ausschließlich als „böse“ beschrieben werden, ist ihre leichte Reizbarkeit legendär, aus der heraus sie oftmals nicht nachvollziehbare Handlungen begehen. Daraus leitet sich nach Meinung vieler die Analogie zum Internet-Troll ab. Für andere ist die Wortherkunft des Troll mit großer Wahrscheinlichkeit keine direkte Übertragung aus der Mythologie. Stattdessen stammen seine Ursprünge wohl von der lautmalerischen Verballhornung des Fischens mit Schleppnetz („trawl“), das als Analogie für das Verhalten eines Trolls gebraucht wurde.



Die allgemeine Weisheit im Umgang mit Internet-Trollen lautet: „Don't feed the trolls!“ („Trolle bitte nicht füttern!“), denn das würde sie nur ermutigen. Das heißt, antworte ihnen nicht, denn das ist die Aufmerksamkeit, nach der sie sich sehnen. Häufig sieht man als Antwort auf einen Troll-Beitrag einfach den Begriff plonk, der das Geräusch des Auftreffens eines Trolls im Killfile des Lesers veranschaulichen soll, d. h. weitere Beiträge des Trolls werden ignoriert.

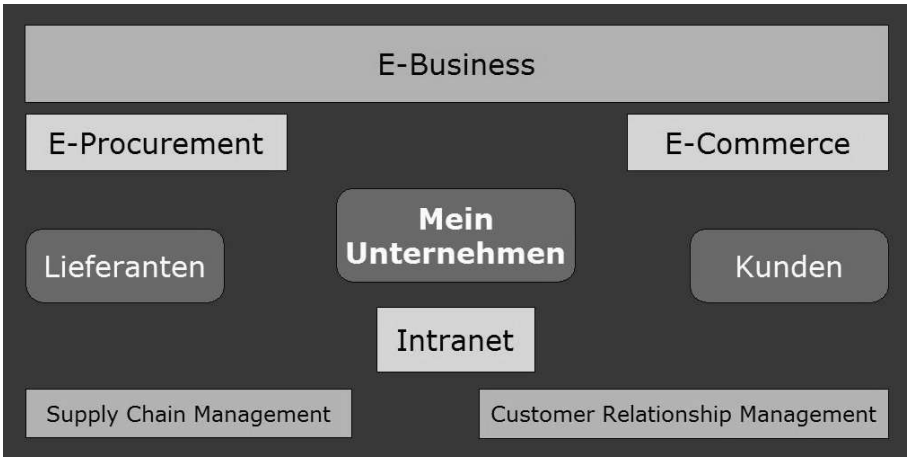
Mitleidige Naturen werfen der armen Kreatur hingegen einen Fisch (z. B. einen Hering) in sein Netz und verstoßen damit gegen die elementare Grundregel des Fütterungsverbotes. Der Ursprung dieser Tradition ist nicht mehr eindeutig nachvollziehbar.

Einerseits soll das zur Verfügung stellen der "Hirnnahrung" Fisch wohl ein Äquivalent zur Bitte "Herr schmeiss Hirn vom Himmel" darstellen und dem Beworfenen zeigen, dass ihm eben jenes fehlt.
Andererseits lassen sich Parallelen zu einer Spielszene aus Monkey Island erkennen, in welcher einem Brückentroll ein roter Hering gegeben werden musste, um von ihm in Ruhe gelassen zu werden.
Aber möglicherweise wird mit dem Fischwurf auch auf die Prügelszenen in Asterix angespielt, die häufiger (in der Umgebung des Fischhändlers Verleihnix) auftreten. Meist trifft der erste Fisch sein Ziel ungewollt, dann beginnt die Dorfschlägerei.

E-Business

Elektronische Geschäftsprozesse bzw. **E-Business** kann sowohl aus der unternehmensinternen Perspektive (Supply-Chain-Management, E-Procurement) als auch aus der unternehmensexternen Perspektive (Elektronische Marktplätze, E-Hubs usw.) gesehen werden.

Eine abgrenzende Begriffsdefinition für **E-Business** ist schwierig zu finden. Der Begriff an sich wurde im Jahre 1998 durch eine Werbekampagne von IBM populär.



Heute versteht man unter E-Business in der Regel alle Methoden der Abwicklung von Geschäften über elektronische Kanäle, wobei das Internet oder zumindest die im Internet verwendeten Techniken und Protokolle eine wesentliche Rolle spielen. Ein wesentlicher Aspekt ist die Aufhebung von Medienbrüchen, wie sie in der konventionelen Geschäftsabwicklung typisch sind. Zudem sollen die Eingriffe von Menschen in den Geschäftsablauf auf das notwendige Minimum reduziert werden. Man spricht in diesem Fall von Straight Through Processing (STP). Dazu ist aber eine weitgehende Integration der Geschäfts-

funktionen erforderlich, die auch unter dem Namen Enterprise Application Integration (EAI) bekannt geworden ist.

Als Vorgängerversion kann der ab 1995 im englischen Sprachraum verwendete Begriff E-Commerce angesehen werden. Nach dem heutigen Begriffsverständnis könnte E-Commerce als Teil des umfassenderen Bereiches E-Business angesehen werden.

Der Oberbegriff E-Business lässt sich nach Art der Teilnehmer kategorisieren in

- B2B Business-To-Business, Unternehmen zu Unternehmen
- B2C Business-To-Consumer, Unternehmen zu Verbraucher
- B2A Business-To-Administration, Unternehmen zu öffentl. Verwaltung
- B2E Business-To-Employee, Unternehmen zu Mitarbeiter
- C2C Consumer-To-Consumer

Logischerweise gibt es auch elektronisch gestützte Geschäftsprozesse von anderen Kategorien, beispielsweise Verbraucher zu öffentlichen Verwaltung in Form der elektronischen Steuererklärung, aber die Abkürzung A2C oder andere sind (noch) nicht allgemein verbreitet.

Volkswirtschaftliche Grundlagen

Die Verständnis für die besonderen Eigenschaften von E-Business erwächst aus der Abkehr von der Neoklassischen Mikroökonomie. Diese setzt unter anderem homogene Güter, vollkommene Markttransparenz und die Abwesenheit von Präferenzen voraus und begibt sich damit auf ein hohes, aber realitätsfernes Abstraktionsniveau. Eine realitätsnähere Beschreibung des Wirtschaftslebens ermöglicht die Neue Institutionenökonomie (NIÖ).

Im Rahmen der Institutionenökonomie spielen die Transaktionskosten eine wichtige Rolle. Das Internet kann Kosten einer Transaktion in der Such- und Anbahnungsphase senken. Auch in der Abwicklungsphase bestehen Chancen zu Senkung der Transaktionskosten. Insgesamt sinken die Kosten für Markttransaktionen und die Koordination über Märkte wird vorteilhafter.

Betriebswirtschaftliche Grundlagen

Ziel einer Umwandlung eines Unternehmens hin zu E-Business ist die Senkung der Kosten im Unternehmen. Maßnahmen um diese Senkung zu erreichen umfassen vor Allem die Optimierung bestehender und Schaffung neuer Geschäftsprozesse auf Basis der Internettechnologien. Durch diese Integration der Geschäftsfunktionen entlang der Wertschöpfungskette soll eine effizientere Geschäftsabwicklung erreicht werden. Durch diese Integration der Geschäftsfunktionen entlang der Wertschöpfungskette soll eine effizientere Geschäftsabwicklung erreicht werden.

LITERATUR

- Merz, Michael: E-Commerce und E-Business: Marktmodelle, Anwendungen und Technologien, 2. Auflage, Heidelberg: dpunkt-Verlag, 2002.
- Thome, Rainer: e-Business, Informatik Spektrum, 18.04.2002, S. 151-153.

E-GOVERNMENT

Das Akronym **E-Government** (aus engl. *Elektronische Regierung*, auch *eGovernment*) steht für *Electronic Government*, zu deutsch etwa "elektronisches Regieren".

Elektronic Government bezeichnet die Nutzung des Internets und anderer elektronischer Medien zur Einbindung der Bürger und Unternehmen in das Verwaltungshandeln sowie zur verwaltungsinternen Zusammenarbeit (Def. nach: Bundesamt für Sicherheit in der Informationstechnik (BSI)). Es ist ein Teilgebiet von E-Business.

E-Government ist damit die Summe aller Verwaltungsprozesse zwischen Behörden („Government“) und Bürgern (G2C), Firmen (G2B) oder anderen Behörden (G2G), die mit Hilfe von Internet-Techniken durchgeführt werden können.

E-Government soll dem Staat vor allem (Personal-)Kosten ersparen, indem die Interaktivität des Internets für den Geschäftsverkehr mit dem Bürger genutzt wird und Personal überwiegend für die Erledigung der eigentlichen Anfrage eingesetzt wird.

Im Rahmen der Umstellung von Verwaltungsprozessen auf E-Government werden üblicherweise alle Aufgaben der Verwaltung in einem Katalog zusammen gefasst und Internetfähigkeit überprüft. Hierbei muss sich die Verwaltung als Anbieter von Produkten sehen, die der Bürger konsumiert.

Einen Rahmen für die Vorgehensweise hat das Bundesamt für Sicherheit in der Informationstechnik (BSI) mit seinem E-Government-Handbuch gesteckt.

Leitprojekt bei eGovernment in Deutschland ist das Portal BUND.DE des Bundesverwaltungsamtes. Koordiniert werden soll eGovernment durch das aus Media@Komm hervor gegangene Portal Deutschland-Online.

Im Rahmen der Initiative BundOnline 2005 hat das Bundesinnenministerium im Jahr 2000 damit begonnen, Standards für die Entwicklung von eGovernement-Systemen zu entwickeln. Ziel war die Vermeidung von kostenintensiven und zeitraubenden Parallel-Entwicklungen, die Sicherstellung einer möglichst großen Kompatibilität der Einzelsysteme und eine möglichst große Unabhängigkeit von den Sortwareherstellern zu gewährleisten. Das Dokument "SAGA - Standards und Architekturen in eGovernment Anwendungen" fasst die wesentlichen Anforderungen zusammen, die grundlegend für eine Förderung durch öffentliche Mittel im Rahmen der Initiative BundOnline 2005 sind.

Ziel ist es, für Bürger und Unternehmen – gleich ob auf kommunaler, regionaler, nationaler, supranationaler oder globaler Ebene – einen umfassenden Lösungsansatz für Interaktionen

mit diesen Behörden zur Verfügung stellen zu können. Eingeschlossen soll der gesamte öffentliche Sektor, bestehend aus Legislative, Exekutive und Judikative sowie öffentlichen Unternehmen werden.

Im Bereich der Judikative hat sich zur Betonung der Unabhängigkeit der Dritten Gewalt der Begriff des E-Justice etabliert.

Weblinks

- http://www.kbst.bund.de/-,54/E-Government.htm - Koordinierungs- und Beratungstelle der Bundesregierung für Informationstechnik in der Bundesverwaltung
- http://www.bund.de/bundonline2005 - Initiative BundOnline 2005
- http://www.deutschland-online.de - Deutschland Online; Koordination eGovernment für Verwaltungen
- http://www.osci.de OSCI - Technischer Standard für eGovernment

Anonymität im Internet

Bei Aktivitäten im Internet fühlen sich viele Benutzer anonym. Diese Anonymität ist jedoch trügerisch. Grundsätzlich erfährt die Gegenseite bei der Kommunikation die IP-Adresse. Doch auch Cookies oder Browserinformationen können ohne Wissen des Anwenders weitergegeben werden.

Mit der IP-Adresse eines Benutzers kann der Anbieter von Internetdiensten die tatsächliche Identität des Benutzers nicht ermitteln, er kann jedoch Hinweise wie den Provider und oft auch noch Land und Region herausfinden. Für die Identität muss eine Anfrage beim Provider erfolgen, dieser besitzt die nötigen Daten. Andere Teilnehmer können sich über das Verhalten dieses Benutzers bei dessen Provider beschweren, welcher dann in der Regel Maßnahmen für diesen Benutzer ergreift (z. B. Sperrung). Strafverfolgungsbehörden können natürlich die Herausgabe der Identität eines Benutzers verlangen, wenn mit dieser IP Straftaten begangen wurden.

Um die IP-Adresse beim Surfen zu verschleiern, werden oft anonymisierende Proxyserver benutzt. Der Proxybetreiber kennt aber immer noch die IP-Adresse des Nutzers, und kann diese auf Anfrage herausgeben. Um das zu vermeiden bauen bestimmte Tools Ketten von Proxies auf, zwischen denen der Verkehr verschlüsselt wird. Diese Variante ist langsam, aber recht sicher, da nur eine fehlende Zwischenstation die Rekonstruktion unmöglich macht. Tools die das verwirklichen sind JAP oder das Hackertool 6/4. Diese sind teilweise diskrediert, also erfüllen ihren Zweck nicht mehr im vollen ursprünglichen Umfang.

Will man anonym Daten veröffentlichen oder Dateien tauschen, kommen anonyme Peer-to-Peer-Netzwerke zum Zug. Sie funktionieren ähnlich, mit mehreren Zwischenstationen und Verschlüsselung an jedem Pfad. Vertreter dieser Sparte sind Freenet, Mnet und Gnunet.

Um anonyme E-Mails versenden zu können oder anonym Usenet-Postings zu erstellen benutzt man so genannte Remailer. Diese funktionieren ähnlich wie ein Proxy, nur für E-Mails: Die Nachricht wird von dem Remailer weiterversendet, so dass dieser als Absender auftaucht. Ein Remailer-Dienst ist Mixmaster.

Strafverfolgungsbehörden haben Schwierigkeiten mit der Aufklärung, wenn bei über das Internet verübten Verbrechen solche Verschleierungsmechanismen genutzt werden. Daher wird von ihrer Seite eine Einschränkung oder sogar eine Illegalisierung solcher Dienste gefordert.

Wichtig ist die Anonymität um die Redefreiheit zu sichern, vorallem in totalitären Staaten, und auch die Privatsphäre vor Rasterdatenerfassung, z.B. durch Werbefirmen zu schützen

Zensur im Internet

Die **Zensur im Internet** unterscheidet sich nicht grundsätzlich von der Zensur anderer Medien.

Ein *Vorzensur* ist aufgrund der dezentralen Struktur des Internets nicht möglich. Die Nachzensur im Internet ist nur schwer, wenn überhaupt zu erreichen. Peer-to-Peer-Projekte wie Freenet erhöhen die Anonymität des Benutzers und auch des Autors.

Der wesentliche Unterschied sind die Staatsgrenzen, die im Internet nicht existieren. Daraus ergibt sich eine hohe Komplexität rechtlicher Fragen, da Unvereinbarkeiten zwischen Rechtssystemen nicht lösbar sind.

Regierungen und staatliche Organe können durch das Sperren von Webseiten, die in ihrem Rechtsbereich liegen, auch die Bürger anderer Staaten von diesen Informationen abhalten, jedoch können sie nicht verhindern, dass die Bürger sich Zugang zu illegalen Informationen verschaffen, die im Ausland liegen.

In Deutschland sind zum Beispiel die Verherrlichung der NS-Kriegsverbrechen oder auch die Verleugnung des Holocaust verboten. Auf US-Servern hingegen können diese Dinge ungestraft verbreitet werden, da sie im US-Recht von der Meinungsfreiheit abgedeckt werden. Ein Vorstoß in Richtung Zensur gab es 2001 vom Bundesland Nordrhein-Westfalen, der aber sowohl technisch als auch rechtlich scheiterte.

China schränkt den Zugriff auf ausländische, vor allem taiwanesische, Nachrichtenseiten ein. Kuba verbietet einen Internetzugriff ohne vorherige Genehmigung.

Aber auch nicht-staatliche Gruppen wie Scientology versuchen, mit juristischen Mitteln kritische Informationen über die Sekte zu unterbinden.

Das US-amerikanische Project for the New American Century, dessen Mitglieder sich in der Regierung Bush (2000-2004) wiederfanden, veröffentlichte im September ein Dokument, in dem dem Internet eine große Bedeutung in der modernen Kriegsführung und Informationspolitik und -beschaffung zukommt.

Auch Suchmaschinen wie Google haben und nutzen die Möglichkeiten zur Zensur. Webseiten, die nicht in einer Suchmaschine aufgeführt werden, können vom Benutzer auch nur schwer gefunden werden (*the creation of 'US Space Forces', to dominate space, and the total control of cyberspace to prevent 'enemies' using the internet against the US*).

Im Februar 2004 zensierte der Internetprovider Freenet.de Webseiten, die sich kritisch zu dem Unternehmen äußerten, indem er einen Teil der Nutzer seines Dienstes, die versuchten, die unternehmenskritischen Seiten aufzurufen, auf andere Webseiten umlenkte. Technisch wurde dies durch einen transparenten Proxy realisiert.

Weblinks

- http://www.ccc.de/censorship/ - Internet-Zensur Informationen vom Chaos Computer Club
- http://www.zensur.freerk.com/index-de.htm - „HOWTO: Internetzensur umgehen“

Appendix

Autoren

Die folgenden 597 Autoren haben an den im WikiReader Internet verwendeten Artikeln mitgearbeitet, ausgelassen sind nicht-angemeldete Benutzer (IP-Adressen):

2501, 3247, 4tilden, AHoerstemeier, Ablaubaer, Adaxl, Aka, Akl, Alex42, AlexR, Alexander.stohr, Ali-Alkohol, Anathema, Andre Engels, Andre Riemann, Andreas B., Andreas Stockter, Andreas kunert, AndreasE, Andrejb, Andrsvoss, Andruil, Angela, Angie, Appius, Ari, Arnd, Arne Koewing, Arne List, ArneBab, ArnoLagrange, Arty, Arved, Asarion, Asb, AssetBurned, Astc, Asteroid7687, Atari-Frosch, Avatar, B, Baroi, Bastic, BeatePaland, Belz, Ben-Zin, Benedikt, Bent, Bernhard55, Betterworld, Bib, Bits-fritz, Bitteloeschen, Black river, Blacklibra, Blaubart, Blubbalutsch, Bmr, Bpascal, Brion VIBBER, Burggraf17, Buxul, CHR, CSonic, Caliga, Capullo, Caramdir, Carter666, Cash walton, Cayman islands, Ce, Ce2, Chd, Christian List, Christian Müller, ChristophLanger, Christopher, Chrizz, Chrono, Cirdan, Cm., CmdrFirewalker, Cnmuc, Coma, CommiM, Coroico, Cpcgm, Crissov, Crux, Csw, D, DWay, DaB., Daboss, Daniel.goehler, Danimilkasahne, DarkBoy, DarkWulf, Darklight, Darkone, Darrn, David Hoeffer, Deever, Deltree, Denisoliver, DeusTron, Devnull, Dg1nsw, Diddi, Diftong, Digitus, Dishayloo, Dlat, Dnaber, Dominik, Dr. Hasenbein, Driessel, Drumknott, Drummerboy, Drzoom, Ducki2, Dumbthingy, Easytouch.at, Echoray, Eckhart Wörner, EdvonSchleck, Eehmke, Eilmeldung, El, El Dirko, Electrocat, Elian, Eloquence, Elvis untot, Elwood j blues, Elya, Emi, Epischel, Erd, ErikDunsing, Ernstl, Erwin E aus U, Euripides, F.Merken, Fab, FabGuy, Faco, Fafer, Fanta, Faraway, Fgb, Filzstift, Finex, Fire, Flofin, Florian, Flups, Focus mankind, Frank12, Freerk, Freibeuter, Fristu, Fritz, Fschoenm, Fuzz, Fxb, GNosis, Gail, Gandi, Geekux, Geof, GeorgGerber, German angst, Geschichtsfan, GianVella, GillianAnderson, Glenn, Glglgl, Gnosis, Gruber, Guido Steenkamp, Guillermo, Gunnar Eberlein, Gurt, Gurumaker, HHK, HaThoRator, Hafenbar, Hagbard, Haicobo, Halli, Hammer1, Hamsta, Harald Spiegel, Harko, Hashar, HaukeZuehl, HbJ, Head, Heiks, Heixe, Hella, Helm, HenrikGebauer, HenrikHolke, Hfastedge, Hfst, Hhdw, Hinrich, HoHun, Hoch auf einem Baum, Hokanomono, Holgernohr, Horst Frank, Hubi, Hutschi, IGEL, Ich, Ich hab hunga, Igelball, Ilja Lorek, Imperator, Interactive, Interbay, Intertorsten, Iqfisch, Irmgard, Irongate, Iuiz, Jacek79, JakobVoss, Jan Niggemann, JanW, Jaybee, Jed, Jekub, Jens Gutzeit, Jensens, Jjanis, John Doe, JohnOwens, Joliver, Jonelo, Josef Spindelböck, Jost ammon, Jsuelwald, Jörny, KaerF, Kai Wasserbäch, Karl Gruber, Karl-Henner, KarstenKasdorf, KarstenSchulz, Katharina, Kdwnv, Keichwa, Keyser Soze, Kfen, Kiker99, Kixx, Kku, Klaus Jesper, Klaus Rilling, Klauseck, Kloeschen, Knarf, Koala, Korre, Koyaanis Qatsi, Kpjas, Kris Kaiser, Kruemelmo, Kubieziel, Kurt Jansson, Langec, Langnickel, Langohr5, Law, Leonard Vertighel, Leonardo, Leseratte, Lhelm, Limasign, Lmh, Lomion, Longamp, LosHawlos, Lothar Kimmeringer, Lukian, Lupino, MAK, MFM, MH, MKelting, MSk, Macmewes, Magic, Magicm247, Magnus, Magnus Manske, Maha, Majonaise, ManfredG, Manu, Marc Tobias Wenzel, MarcelJanus, Marcela, Marco Krohn, Marioj, Markus Schweiß, MarkusWinkler, Marlowe, Marti7D3, Martin Aggel, Martin Götzer, Martin.k, Martinroell, Mastad, MasterLR, Mathias Schindler, Matthias, Matthäus Wander, Matusz, MauriceKA, Maus-Trauden, Maveric149, Maxberger, Media lib, Medusa, Memowe, Meph666, Michael Hasin, Michael Schubart, Michael Zeilfelder, Michael w, Michael.chlistalla, MichaelB., MichaelDiederich, MichaelKoch, MikeTheGuru, Mikegr, Mikenolte, Mikl, Mikue, Mirer, Momomu, Montauk, Moolsan, Morget, MrFixit, MrTux, Ms1203, Mulk, Muns, Mwka, Mxr, Myr, Myrisa, N-true, Nainoa, Nankea, Nasir, Nd, Nephelin, Nerd, NetReaper, Netguru, Nevid, NewPapillon, Nick-less, Nikai, Ninjamask, Nocturne, Nonanet, Norbert, Norri, Nx7000, OWeh, OderWat, Odin, Ohohlfeld, Oktaeder, Olemo, OliD, Orbiter, Orkenspalter, Ostrock, Ottsch, Owltom, Paddy, Pal05, Pandaemonium, Parbit, Parka Lewis, Pascal Auricht, PatriceNeff, Patrick Hanft, Patruh, Paul Ebermann, PeFu, PeKron, Peacemaker, Peregrine981, Perlentaucher, PeterBonn, Petit, Petkli, Philipp-Weissenbacher, Phlow.net, Phulab, Pietz, Pikachu, Piko, Pinguin.tk, Pit, Pkn, Plasmagunman, Plenz, Poldi, Priwo, Proggy, Progman, Psypath, Pyrdracon, Qpaly, RKraasch, Rainer Wasserfuhr, RalfZosel, Rapunzel, Raven, Raymond, Redumo, Remi, Rho, Riptor, Rivi, Robb, RobbyBer, Robert G. Siebeck, RobertLechner,

Rolf Weirauch, RolfM, RolfS, Rolz-reus, Rrdd, SMike, STephan Kambor, Sadduk, Sansculotte, Sbeyer, Schattenraum, Schewek, Schnargel, Schoos, Schrottvogel, Sebastian, Sebastian Hagedorn, SebastianBreier, SebastianWilken, Seef, Sei, Sesalo, Sevenstar, Siggi sorglos, Sikilai, SilentSurfer, Simeon Kienzle, Skram, Skriptor, SkyJumpy, Smurf, Snc, Softeis, Sonium, Southpark, Sovok, Spirou44, Srbauer, Stefan Kühn, Stefan Ruehrup, StefanRybo, Stefanwege, Steffen, SteffenB, Stern, Stoph, Str, Strangemeister, Stuffi, Stw, Suesan, Sveith, Swacker, Swing, Synapse, Synthetik, Systemdefender, Tabacha, Tadzio, Tali, Tamino, Taprogge, Terabyte, Thdoerfler, The weaver, The-Me, TheK, Theclaw, Theyak, Thmueller, Thom, Thomas Fernstein, Thomas W., Thomas.gutbier, ThomasGigold, Three of Five, Tieno, Till, Tilmanb, Tilo, Tim Pritlove, Tj, Tkarcher, Tobias Conradi, Tobo, Toka, TomK32, TorPedo, TorsTen, Torsten Henschel, Torsten.otto, TorstenNetzel, Transparent, Traroth, Triebtäter, Triton, Trixium, Trugbild, Tsor, Tsr, Tufkas, TuxBender, Tuxerado, Uebs, Uliuli, Ulrich Rosemeyer, Ulrich.fuchs, UncleOwen, Unukorno, Urbanus, Urs, UweFriedrich, Uweschwoebel, Vanlenderius, VerwaisterArtikel, Vinci, Vmk, Vulture, W ds, WKr, Waelder, Walter, Walter Koch, Warp, Wb2000, Webdream, Webkid, Weede, WeißNix, Whotithi, Wiegand, Wikinator, Wikipedia ce, Wikipedianer, Wilinckx, Windy, Wing, Winzkling, Wipape, Wiska Bodo, Wolfgang1018, Woodworker, Wst, Wurblzap, XTaran, Xell, Xeospeed, Xtian, Yas, Youandme, ZaphodB, Zascha, Zebbo, ZelleAP, Zenogantner, Zenon, Zeus™

QUELLENVERZEICHNIS

ARPANET um 20:49, 16. Apr 2004
Akustikkoppler um 19:56, 16. Apr 2004
Al_Gore um 11:34, 7. Mai 2004
Anonymität_im_Internet um
Auszeichnungssprache um 00:43, 21. Apr 2004
Backdoor um 12:48, 17. Apr 2004
Barrierefreies_Internet um 18:47, 3. Mai 2004
BitTorrent um 00:37, 30. Apr 2004
Bustopologie um 23:20, 27. Apr 2004
Chaos_Computer_Club um 21:54, 4. Mai 2004
Chat um 21:00, 11. Mai 2004
Client-Server-System um 16:27, 21. Dez 2003
Common_Gateway_Interface um 00:41, 27. Apr 2004
Computernetzwerk um 19:23, 10. Mai 2004
Computersicherheit um 16:14, 6. Apr 2004
Computervirus um 13:21, 13. Mai 2004
Computerwurm um 20:38, 8. Mai 2004
Cookie um 14:03, 18. Apr 2004
Cracker um 10:10, 22. Apr 2004
DENIC um 20:43, 27. Apr 2004
DMZ um 23:01, 26. Apr 2004
Denial_of_Service um 22:10, 12. Mai 2004
Dialer um 11:21, 5. Mai 2004
Digital_Subscriber_Line um 09:37, 12. Mai 2004
Domain_Name_System um 08:56, 12. Mai 2004
Domäne_(Internet) um 09:21, 14. Apr 2004
Download um 17:37, 28. Apr 2004
E-Business um 09:23, 24. Apr 2004
E-Government um 11:00, 27. Apr 2004
E-Mail um 12:35, 8. Mai 2004
Emoticon um 18:03, 6. Mai 2004
Englische_Sprache_im_Internet: 11:46, 9. Mai 2004
Eric_Allman um 14:57, 29. Apr 2004
Extensible_Markup_Language: 18:41, 10. Mai 2004
File_Transfer_Protocol um 09:29, 13. Mai 2004
Firewall um 16:35, 10. Mai 2004
GNU_Privacy_Guard um 16:14, 11. Mai 2004
Gopher um 14:16, 7. Mai 2004
Hyperlink um 23:43, 11. Mai 2004
Hypertext um 23:13, 1. Mai 2004
Hypertext_Markup_Language: 19:55, 4. Mai 2004
Hypertext_Transfer_Protocol; 23:43, 29. Apr 2004
I-mode um 13:21, 13. Mai 2004
IP-Spoofing um 13:58, 19. Feb 2004
IP-Telefonie um 22:38, 10. Mai 2004
IPsec um 17:03, 10. Mai 2004
IPv4 um 12:18, 10. Mai 2004
IPv6 um 22:46, 1. Mai 2004
Instant_Messaging um 11:24, 4. Mai 2004
Integrated_Services_Digital_Network um 11:22, 8. Mai 2004
Internet um 23:50, 1. Mai 2004
Internet-Protokoll-Familie um 23:13, 29. Apr 2004
Internet2 um 18:58, 20. Apr 2004
Internet_Corporation_for_Assigned_Names_and_Numbers um 18:37, 12. Apr 2004
Internet_Engineering_Task_Force um 23:42, 1. Mai 2004
Internet_Protocol um 23:50, 1. Mai 2004
Internet_Relay_Chat um 13:51, 10. Mai 2004
Internetarchiv um 23:43, 1. Mai 2004
Internetdienstanbieter um 18:00, 23. Apr 2004
Internetsucht um 03:52, 12. Mai 2004
Java-Applet um 23:20, 11. Mai 2004
Jonathan_Postel um 00:16, 5. Mär 2004

Klammeraffe um 22:52, 28. Apr 2004
LAMP um 19:21, 14. Apr 2004
Lesezeichen_(Internet) um 06:54, 29. Apr 2004
Mailbox_(Computer) um 13:45, 8. Mai 2004
Mailingliste um 22:42, 22. Apr 2004
Modem um 20:08, 27. Apr 2004
Mozilla um 17:42, 10. Mai 2004
Napster um 20:35, 24. Mär 2004
Netiquette um 21:24, 28. Apr 2004
Network_Address_Translation um 12:14, 9. Mai 2004
Netzkultur um 01:48, 21. Apr 2004
Netzwerkprotokoll um 10:17, 23. Apr 2004
OSI-Modell um 20:17, 10. Mai 2004
PHP um 21:19, 11. Mai 2004
PPP_over_Ethernet um 22:31, 16. Apr 2004
Peer-to-Peer um 22:54, 11. Mai 2004
Ping_(Datenübertragung) um 20:33, 20. Mär 2004
Point-to-Point_Protocol um 11:20, 26. Apr 2004
Proxy um 20:22, 27. Apr 2004
Request_for_Comments um 04:55, 5. Mai 2004
Ringtopologie um 23:52, 27. Apr 2004
Robert_E._Kahn um 21:43, 13. Apr 2004
Root-Server um 00:20, 29. Apr 2004
Rootkit um 19:52, 27. Apr 2004
Routing um 11:59, 8. Mai 2004
Secure_Shell um 00:18, 24. Apr 2004
Security_through_Obscurity um 19:06, 7. Apr 2004
Spam um 16:27, 10. Mai 2004
Spyware um 08:48, 3. Apr 2004
Standleitung um 11:11, 8. Mai 2004
Sterntopologie um 11:29, 3. Mai 2004
Subdomain um 07:04, 29. Apr 2004
Suchmaschine um 09:40, 10. Mai 2004
TOFU um 19:39, 30. Apr 2004
Tauschbörse um 16:28, 12. Mai 2004
Telnet um 09:15, 27. Apr 2004
Tim_Berners-Lee um 20:26, 24. Apr 2004
Top_Level_Domain um 10:43, 13. Mai 2004
Traceroute um 20:04, 27. Apr 2004
Transmission_Control_Protocol um 09:40, 29. Apr 2004
Trojanisches_Pferd_(Computerprogramm) um 07:40, 30. Apr 2004
Troll_(Internet) um 16:35, 4. Mai 2004
Uniform_Resource_Identifier um 10:09, 22. Apr 2004
Universal_Mobile_Telecommunications_System um 21:52, 12. Mai 2004
Upload um 20:01, 29. Apr 2004
Usenet um 23:43, 10. Mai 2004
User_Datagram_Protocol um 16:48, 10. Mai 2004
VCard um 12:00, 1. Mär 2004
Vermaschtes_Netzwerk um 09:20, 28. Apr 2004
Vinton_G._Cerf um 11:44, 7. Mai 2004
Web-Bug um 17:48, 6. Apr 2004
Webbrowser um 17:55, 12. Mai 2004
Webdesign um 17:20, 25. Apr 2004
Webforum um 11:01, 1. Mai 2004
Weblog um 13:53, 10. Mai 2004
Webportal um 01:47, 29. Apr 2004
Webseite um 17:34, 7. Mai 2004
Webserver um 23:30, 30. Mär 2004
Wiki um 20:50, 10. Mai 2004
Wireless_Application_Protocol um 12:58, 10. Apr 2004
Wireless_LAN um 21:50, 12. Mai 2004
World_Wide_Web um 23:42, 3. Mai 2004
World_Wide_Web_Consortium um 15:28, 29. Apr 2004

GNU Freie Dokumentationen Lizenz

This is an unofficial translation of the GNU Free Documentation License into German. It was not published by the Free Software Foundation, and does not legally state the distribution terms for documentation that uses the GNU FDL–only the original English text of the GNU FDL does that. However, we hope that this translation will help German speakers understand the GNU FDL better.

Dies ist eine inoffzielle deutsche Übersetzung der GNU Free Documentation License. Sie ist nicht von der Free Software Foundation herausgegeben und erläutert nicht die Bedingungen der GNU FDL – Dies tut nur der original englische Text der GNU FDL. Dennoch hoffen wir, dass diese Übersetzung mit dazu beiträgt deutschsprachigen Personen das Verstehen der GNU FDL zu erleichtern.

Präambel

Der Zweck dieser Lizenz ist es, ein Handbuch, Textbuch oder ein anderes zweckdienliches und nützliches Dokument frei, im Sinne von Freiheit, zu machen; jedermann die Freiheit zu sichern, es zu kopieren und mit oder ohne Änderungen daran, sowohl kommerziell als auch nicht kommerziell weiter zu verbreiten.
Weiterhin sichert diese Lizenz einem Autor oder Verleger die Möglichkeit, Anerkennung für seine Arbeit zu erhalten ohne für Änderungen durch Andere verantwortlich gemacht zu werden.

Diese Lizenz ist eine Art des „copyleft“, was bedeutet, dass von diesem Dokument abgeleitete Werke ihrerseits in derselben Weise frei sein müssen.
Dies vervollständigt die GNU General Public License, die eine „copyleft“-Lizenz ist, und für freie Software entworfen wurde.

Diese Lizenz wurde für Handbücher für freie Software entworfen, denn frei Software braucht freie Dokumentation: Ein freies Programm sollte von Handbüchern begleitet sein, die dieselben Freiheiten bieten, die auch die Software selbst bietet.
Diese Lizenz ist aber nicht auf Softwarehandbücher beschränkt; vielmehr kann sie für jede Art von textuellen Werken verwendet werden, unabhängig davon, was das Thema ist, oder ob es als gedrucktes Buch veröffentlicht wurde. Wir empfehlen diese Lizenz prinzipiell für Werke, die als Anleitungen oder Referenzen dienen sollen.

1. Anwendbarkeit und Definitionen

Diese Lizenz findet Anwendung auf jedes Handbuch oder andere Werk, unabhängig von dem Medium, auf dem es erscheint, das einen vom Rechteinhaber eingefügten Hinweis enthält, der besagt, dass das Werk unter den Bedingungen dieser Lizenz verbreitet werden darf.
Ein solcher Hinweis gewährt eine weltweit gültige, tantiemenfreie und zeitlich unbefristete Lizenz, die es gestattet das Werk, unter den hier festgelegten Bedingungen, zu nutzen.
Der Begriff Dokument wird im Folgenden für alle solche Handbücher und Werke verwendet.
Jede Person kann Lizenznehmer sein und wird im Folgenden mit Sie angesprochen.
Sie akzeptieren diese Lizenz, wenn Sie ein Dokument derart kopieren, verändern oder verteilen, dass Sie gemäß den Gesetzen zum Copyright die Erlaubnis benötigen.

Eine modifizierte Version des Dokumentes steht für jedes Werk, das das Dokument als Ganzes oder in Teilen enthält, sowohl auf Datenträger kopiert, als auch mit Änderungen und/oder in andere Sprachen übersetzt.

Ein zweitrangiger Abschnitt ist ein benannter Anhang oder eine Enleitung des Dokumentes, der sich ausschließlich mit dem Verhältnis des Autors oder Verlegers des Dokumentes zu dem eigentlichen Thema des Dokumentes (oder damit zusammenhängender Dinge) beschäftigt, und der nichts enthält, das direkt zu dem eigentlichen Thema gehört. (Wenn das Dokument beispielweise ein Buch über Mathematik ist, dann darf ein zweitrangiger Abschnitt nichts über Mathematik enthalten).
Dies kann eine historische Beziehung zu dem Thema, oder damit zusammenhängender Dinge, oder von gesetzlicher, gesellschaftlicher, philosophischer, ethischer oder politischer Art sein, die das Thema betreffen.

Die unveränderlichen Abschnitte sind benannte zweitrangige Abschnitte, deren Titel als unveränderlicher Abschnitt in dem Lizenhinweis, der das Dokument unter diese Lizenz stellt, aufgeführt sind.
Wenn ein Abschnitt nicht in die oben stehende Definition eines zweitrangigen Abschnittes passt, dann ist es nicht erlaubt diesen Bereich als unveränderlichen Bereich zu kennzeichnen.

Umschlagtexte sind bestimmte, kurze Textstücke, die als vorderer Umschlagtext oder als hinterer Umschlagtext in der Notiz benannt werden, die besagt, dass das Dokument unter dieser Lizenz freigegeben ist.
Ein vorderer Umschlagtext kann bis zu 5 Worte enthalten, ein hinterer Umschlagtext bis zu 25 Worte.

Eine transparente Kopie des Dokumentes bezeichnet eine maschinenlesbare Kopie, dargestellt in einem Format, dessen Spezifikationen allgemein verfügbar sind, und das geeignet ist das Dokument auf einfache Weise mit einem allgemeinen Texteditor oder (für Bilder, die aus Pixeln bestehen) mit einem allgemeinen Bildberabeitungsprogramm oder (für Zeichnungen) mit einem häufig verfügbaren Zeichenprogramm zu überarbeiten, und das geeignet ist es als Eingabe für Textformatierer zu verwenden, oder als Eingabe für automatische Konvertierungsprogramme, die eine Reihe von unterschiedlichen Formaten erzeugen, die ihrerseits als Eingabe für Textformatierer verwendet werden können. Eine Kopie in ein anderes transparentes Dateiformat dessen Auszeichnung oder das fehlen der Auszeichnungen derart beschaffen sind, nachfolgende Modifikationen durch die Leser zu verhindern oder zu erschweren ist nicht transparent
Ein Bildformat ist nicht transparent, wenn es für eine wesentliche Menge von Text verwendet wird.
Eine Kopie, die nicht transparent ist, wird als opak bezeichnet.

Beispiele verwendbarer Formate für transparente Kopien schliessen einfachen ASCII-Text ohne Auszeichnungen, TeX-info Eingabe, LaTeX-Eingabeformat, SGML oder XML, sofern die verwendete DTD öffentlich verfügbar ist, sowie standardkonformes, einfaches HTML, Postscript oder PDF, die für Veränderungen durch Menschen entworfen sind, ein.

Beispiele für transparente Bildformate sind unter anderem PNG, XCF und JPG.
Opake Formate sind unter anderen solche proprietären Formate, die nur von proprietären Textverarbeitungsprogramm gelesen und bearbeitet werden können, SGML oder XML deren DTD und/oder Verarbeitungswerkzeuge nicht allgemein verfügbar sind, und maschinengeneriertes HTML, PostScript oder PDF, das von manchen Textverarbeitungsprogrammen nur zu Ausgabezwecken erzeugt wird.

Mit Titelseite wird in einem gedruckten Buch die eigentliche Titelseite sowie die direkt darauf folgenden Seiten bezeichnet, die all das in lesbarer Form enthalten, was in dieser Lizenz gefordert ist, dass es auf der Titelseite erscheinen muss.
Für Werke, die in Formaten vorliegen, die keine Titelseiten haben, gilt als Titelseite der Text, der der auffälligsten Darstellung des Titels des Werkes direkt folgt, aber noch vor dem Inhalt des Werkes steht.

Ein Abschnitt mit dem Titel xyz bezeichnent einen benannten Unterbereich des Dokumentes, dessen Titel entweder genau xyz ist, oder der xyz in Anführungszeichen enthält, der einem Text folgt, der xyz in eine andere Sprache übersetzt. (Hier steht xyz für einen speziellen Abschnittsnamen, der im Folgenden erwähnt wird wie „Danksagung“ (Acknowledgements), „Widmung“ (Dedications), „Anmerkung“ (Endorsement) oder „Historie“ (History).).
Den Titel erhalten eines Abschnittes bedeutet, dass beim Modifizieren des Dokumentes dieser Abschnitt mit dem Titel xyz bleibt, wie es in dieser Definition festgelegt ist.

Das Dokument kann direkt hinter der Notiz, die besagt, dass das Dokument unter dieser Lizenz freigegeben ist, Garantieausschlüsse enthalten. Diese Garantieausschlüsse werden so behandelt, asl seien sie als Referenzen in diese Lizenz eingeschlossen, allerdings nur um Garantien auszuschliessen: Jede andere Implizierung, die dieser Ausschluss hat ist ungültig und keine Wirkung im Sinne dieser Lizenz.

2. Datenträgerkopien

Sie dürfen das Dokument auf jedem Medium sowohl kommerziell als auch nicht kommerziell kopieren und verbreiten, vorausgesetzt, dass diese Lizenz, die Copyright-Hinweise sowie der Lizenzhinweis, der besagt, dass diese Lizenz auf das Dokument anzuwenden ist, in allen Kopien reproduziert wird, und dass keine weiteren Bedingungen jeglicher Art zu denen dieser Lizenz hinzugefügt werden.
Sie dürfen in den Kopien, die Sie erstellen oder verbreiten, keinerlei technische Maßnahmen treffen um das Lesen oder das weitere Kopieren zu erschweren oder zu kontrollieren. Dennoch dürfen Sie Gegenleistungen für Kopien akzeptieren. Wenn Sie eine ausreichend große Menge von Kopien verteilen, müssen Sie zusätzlich die bestimmungen von Ziffer 3 beachten.
Sie können ausserdem unter denselben Bedingungen, die oben angeführt sind, Kopien verleihen und sie können Kopien auch öffentlich bewerben.

3. Kopien in Stückzahlen

Wenn Sie gedruckte Kopien des Dokumentes (oder Kopien auf Medien, die üblicherweise gedruckte Umschläge haben), in einer Stückzahl von mehr als 100 veröffentlichen, und der Lizenzhinweis des Dokumentes Umschlagtexte verlangt, müssen die Kopien in Hüllen verpackt sein, die alle diese Umschlagtexte klar und lesbar enthalten. Die vorderen Umschlagtexte auf dem vorderen Umschlag, die hinteren Umschlagtexte auf dem hinteren Umschlag.
Beide Umschläge müssen Sie ausserdem klar und lesbar als den Herausgeber dieser Kopien benennen.
Der vordere Umschlag muss den gesamten Titel darstellen, mit allen Worten gleich auffällig und sichtbar. Sie können weiteres Material den Umschlägen hinzufügen.
Das Kopieren mit Änderungen, die auf Umschläge begrenzt sind, können, so lange der Titel des Dokuments erhalten bleibt, ansonsten als Datenträgerkopien behandelt werden.

Wenn der vorgeschriebene Text für einen der Umschläge zu umfangreich ist um lesbar zu bleiben, sollten Sie den ersten der aufgelisteten Texte auf den aktuellen Umschlag nehmen (so viel wie vernünftigerweise möglich ist) und den Rest auf direkt angrenzenden Seiten.

Wenn Sie mehr als 100 opake Kopien veröffentlichen oder verbreiten, müssen Sie entweder eine maschinenlesbare, transparente Kopie jeder opaken Kopie beilegen, oder mit bzw. in jeder opaken Kopie eine Computer-Netzwerk Adresse angeben, von wo die allgemeine, netzwerk benutzende Öffentlichkeit, Zugriff zum Download einer kompletten transparenten Kopie über öffentliche Standardnetzwerkprotokolle hat.

Wenn Sie sich für die letztere Möglichkeit entscheiden, müssen Sie mit Beginn der Verbreitung der opaken Kopien in Stückzahlen, zumutbare und vernünftige Schritte unternehmen, um sicher zu stellen, dass die transparenten Kopien mindestens ein Jahr nach der Auslieferung der letzten opaken Kopie (direkt oder über einen Agenten oder Händler) dieser Ausgabe an die Öffentlichkeit, an der genannten Adresse verfügbar bleiben.

Es ist erbeten, aber nicht gefordert, dass Sie ausreichend lange vor der Auslieferung einer grösseren Menge von Kopien, Kontakt mit den Autoren des Dokumentes aufnehmen, um jenen die Möglichkeit zu geben, Ihnen eine aktualisierte Version des Dokumentes zuzuleiten.

4. Modifikationen

Unter den obigen Bedingungen unter Ziffer 2 und 3 können Sie modifizierte Versionen kopieren und verbreiten, vorausgesetzt, dass Sie die modifizierte Version unter exakt dieser Lizenz herausgeben, wobei die modifizierte Version die Rolle des Dokumentes einnimmt, und dadurch die weitere Modifikation und Verbreitung an jeden Lizensieren, der eine Kopie davon besitzt.

Zusätzlich müssen Sie die folgenden Dinge in der modifizierten Version beachten:

1. Benutzen Sie auf der Titelseite (und auf Umschlägen, sofern vorhanden) einen Titel, der sich von dem Titel des Dokumentes und von früheren Versionen unterscheidet. (Die früheren Versionen sollten, wenn es welche gibt, in dem Abschnitt Historie aufgelistet werden.)
   Sie können denselben Titel wie den einer Vorgängerversion verwenden, wenn der ursprüngliche Herausgeber damit einverstanden ist.
1. Geben Sie auf der Titelseite eine oder mehrere Personen oder Einheiten, die als Autoren auftreten können, als für die Modifikationen verantwortliche Autoren der modifizierten Version, zusammen mit mindestens fünf der ursprünglichen Autoren der Ursprungsversion an (alle vorherige Autoren, wenn es weniger als fünf sind), es sei denn diese befreien Sie von dieser Notwendigkeit.
2. Geben Sie auf der Titelseite den Namen des Herausgebers als Herausgeber an.
3. Erhalten Sie alle Copyright-Vermerke des Dokumentes.
4. Setzen Sie einen passenden Copyright-Vermerk für Ihre Modifikationen direkt hinter die anderen Copyright-Vermerke.
5. Schliessen Sie direkt hinter den Copyright-Vermerken einen Lizenzhinweis ein, der die öffentliche Erlaubnis erteilt, die modifizierte Version unter den Bedingungen dieser Lizenz zu benutzen, wie es im Anhang weiter unten beschrieben ist.
6. Erhalten Sie im Copyright-Vermerk die komplette Liste der unveränderlichen Abschnitte und obligatorischen Umschlagtexte, die in dem Lizenzvermerk des Dokumentes aufgeführt sind.
7. Schliessen Sie eine unveränderte Kopie dieser Lizenz mit ein.
8. Erhalten Sie den Abschnitt „Historie“. Erhalten Sie den Titel und fügen Sie einen Punkt hinzu der mindestens den Titel, das Jahr, die neuen Autoren und Herausgeber, wie sie auf der Titelseite aufgeführt sind, enthält. Sollte es keinen Abschnitt Historie geben, dann erstellen Sie einen, der Titel, Jahr, Autor und Herausgeber des Dokumentes, wie auf der Titelseite angegeben, enthält und fügen Sie einen Punkt hinzu, der die modifizierte Version wie oben dargestellt beschreibt.
9. Erhalten Sie die Netzwerkadresse, die angegeben wurde, um Zugang zu einer transparenten Kopie zu gewähren, sowie entsprechend angegebene Adressen früherer Versionen, auf denen das Dokument aufbaute. Diese Angaben können in den Abschnitt Historie verschoben werden. Sie können die Netzwerkadresse weglassen, wenn sie sich auf ein Werk bezieht, das mindestens 4 Jahre vor dem Dokument selbst veröffentlicht wurde, oder wenn der ursprüngliche Herausgeber der Version, auf die sich die Adresse bezieht, seine Erlaubnis erteilt.
10. Erhalten Sie für alle Abschnitt, die als Danksagungen(Acknowledgements) oder Widmungen(Dedications) überschrieben sind, den Titel sowie die Substanz und den Ton aller vom Geber gemachten Danksagungen und/oder Widmungen in diesem Abschnitt.
11. Erhalten Sie alle unveränderlichen Abschnitte unverändert, sowohl im Titel als auch im Text. Abschnittsnummern oder dergleichen gelten hierbei nicht als Teil des Titels.
12. Löschen Sie alle Abschnitte, die als Anmerkungen(Endorsements) überschrieben sind. Ein solchen Abschnitt sollte nicht in der modifizierten Version enthalten sein.
13. Benennen Sie keinen Abschnitt in Anmerkungen um, oder in einen Namen, der in Konflikt mit einem unveränderlichen Abschnitt gerät.
14. Erhalten Sie alle Garantieausschlüsse.

Wenn die modifizierte Version neue Vorspannabschnitte oder Anhänge enthält, die zweitrangige Abschnitte sein können, und die kein vom Dokument kopiertes Material enthalten, können Sie, nach Ihrem Belieben, einige oder alle diese Abschnitte als unveränderliche Abschnitte in die Lizenzanmerkung der modifizierten Version aufnehmen. Diese Titel müssen sich von allen anderen Titeln unterscheiden.

Sie können einen Abschnitt Anmerkungen anfügen, sofern dieser nichts als Bemerkungen, verschiedener Stellen, zu der modifizierten Version enthält.

Beispielsweise Publikumsreaktionen oder eine Mitteilung, dass der Text von einer Organisation als maßgebliche Definition eines Standards geprüft wurde.

Sie können einen Teil mit bis zu fünf Worten als vorderen Umschlagtext und einen mit bis zu 25 Worten als hinteren Umschlagtext an das Ende der Liste mit den Umschlagtexten der modifizierten Version hinzufügen.
Nur je ein Teil für den vorderen Umschlagtext und den hinteren Umschlagtext können von jeder Einheit hinzugefügt (oder durch entsprechende Anordnung erstellt) werden.

Wenn das Dokument bereits einen Umschlagtext für denselben Umschlag enthält, das von Ihnen oder der Einheit, in deren Namen Sie tätig sind, bereits früher eingefügt wurde, dürfen Sie keine neue hinzufügen. Sie können aber den alten ersetzen, wenn sie die ausdrückliche Genehmigung des Herausgebers haben, der den früheren Text eingefügt hat.

Der/die Autor(en) und Herausgeber des Dokumentes geben duch diese Lizenz weder implizit noch explizit die Erlaubnis ihren Namen für Werbung in den Anmerkungen der modifizierten Version zu benutzen.

5. Dokumente Kombinieren

Sie können mehrere Dokumente, die unter dieser Lizenz freigegeben sind, unter den Bedingungen unter Ziffer 4 für modifizierte Versionen miteinander kombinieren, vorausgesetzt, dass in der Kombination alle unveränderlichen Abschnitte aller Originaldokumente, enthalten sind, und dass Sie diese alle in der Liste der unveränderlichen Abschnitte der Lizenzanmerkung des kombinierten Dokumentes aufführen, sowie alle Garantieausschlüsse erhalten.

Das kombinierte Werk braucht nur eine Kopie dieser Lizenz zu enthalten, und mehrere identische unveränderliche Abschnitte können durch eine einzelne Kopie ersetzt werden.
Wenn es mehrere unveränderliche Abschnitte mit unterschiedlichem Inhalt aber gleichem Namen gibt, machen Sie den Namen eindeutig, indem Sie am Ende des Titels, in Anführungszeichen, den Namen des original Autors oder Herausgebers, falls bekannt, oder andernfalls eine eindeutige Nummer anhängen.
Machen Sie dasselbe mit den Titeln der Abschnitte in der Liste der unveränderlichen Abschnitte im Lizenzhinweis des kombinierten Werkes.

In der Kombination müssen Sie alle Abschnitte mit dem Titel Historie in den unterschiedlichen Dokumenten zu einem einzelnen Abschnit Historie zusammenführen; entsprechend verfahren Sie mit den Abschnitten Danksagungen und Widmungen. Sie müssen alle Abschnitte mit dem Titel Anmerkungen löschen.

6. Sammlungen von Dokumenten

Sie können eine Sammlung von Dokumenten erstellen, bestehend aus diesem Dokument und weiteren, unter dieser Lizenz stehenden Dokumenten, wobei Sie die einzelnen Kopien dieser Lizenz in den verschiedenen Dokumenten durch eine einzelne Kopie, die in der Sammlung enthalten ist, ersetzen, vorausgesetzt, Sie befolgen in allen andern Punkten, für jedes der Dokumente, die Regeln für Datenträgerkopien.

Sie können ein einzelnes Dokument aus einer solchen Sammlung herausziehen und einzeln unter dieser Lizenz verbreiten, vorausgesetzt, Sie fügen eine Kopie dieser Lizenz in das extrahierte Dokument ein, und befolgen ansonsten die Bedingungen dieser Lizenz für Datenträgerkopien.

7. Aggregation mit unabhängigen Werken

Eine Zusammenstellung des Werkes, oder von Ableitungen davon, mit anderen, separaten und unabhängigen Dokumenten oder Werken, in oder auf demselben Band eines Speicher- oder Verbreitungsmediums, wird dann eine Aggregation genannt, wenn die Copyrights der Zusammenstellung nicht dazu verwendet werden die Rechte der Benutzer, die für die einzelnen Werke gewährt werden, stärker zu beschränken als dies durch die Lizenzen der einzelnen Werke geschieht.
Wenn das Werk in einer Aggregation vorhanden ist, so gilt diese Lizenz nicht für die anderen Werke dieser Aggregation, die keine Ableitung des Dokumentes sind.

Wenn die Bestimmungen für die Umschlagtexte aus Ziffer 3 Anwendung finden, und wenn das Dokument weniger als die Hälfte der gesammten Aggregation ausmacht, dann können die Umschlagtexte auf Seiten gesetzt werden, die das Dokument innerhalb der Aggregation umschliessen, oder auf das elektronische Äquivalent eines Umschlages, wenn das Dokument in elektronischer Form vorliegt.
Andernfalls müssen sie auf gedruckten Umschlägen erscheinen, die das gesamte Werk umschliessen.

8. Übersetzung

Übersetzungen werden als eine Art von Modifikationen betrachtet. Damit können Sie eine Übersetzung des Dokumentes unter den Bestimmungen von Ziffer 4 verbreiten.
Um die unveränderlichen Abschnitte durch eine Übersetzung zu ersetzen, benötigen Sie die spezielle Erlaubnis des Copyright-Inhabers. Sie können allerdings Übersetzungen von einigen oder allen unveränderlichen Abschnitten zu den original Versionen der unveränderlichen Abschnitte hinzufügen.
Sie können eine Übersetzung dieser Lizenz und allen Lizenzhinweisen im Dokument sowie allen Garantieausschlüssen hinzufügen, vorausgesetzt, dass Sie ebenso die originale englische Version dieser Lizenz und aller Hinweise und Ausschlüsse beifügen.
Sollten die Übersetzung und die Originalversion dieser Lizenz oder eines Hinweises oder Ausschlusses voneinander abweichen, so hat die Originalversion vorrang.

Wenn ein Abschnitt des Dokumentes als Danksagung, Widmungen oder Historie überschrieben ist, so erfordert die Forderung (Ziffer 4) den Titel dieses Abschnittes zuerhalten, die Änderung des aktuellen Titels.

9. Abschlussbestimmungen

Sie dürfen dieses Dokument nicht kopieren, verändern, unterlizensieren oder verteilen mit der Ausnahme, dass Sie es ausdrücklich unter dieser Lizenz tun.
Jedweder andere Versuch zu kopieren, zu modifizieren, unter zu lizensieren oder zu verbreiten ist unzulässig und führt automatisch zum Entzug der durch diese Lizenz gewährten Rechte. Dennoch verlieren jene Parteien, die von ihnen Kopien oder Rechte unter dieser Lizen erhalten haben, nicht Ihre Rechte, so lange sie sich in völliger Übereinstimmung mit der Lizenz befinden.

10. Spätere Überarbeitungen dieser Lizenz

Die Free Software Foundation kann von Zeit zu Zeit neue, überarbeitete Versionen der GNU Free Dokumentation License veröffentlichen. Diese neuen Versionen werden im Geiste gleich bleiben, können sich aber in Details unterscheiden um neuen Problemen oder Besorgnissen gerecht zu werden.
Siehe: http://www.gnu.org/copyleft/
Jede Version dieser Lizenz erhält eine eigene Versionsnummer.
Wenn das Dokument bestimmt, dass eine bestimmt numerierte Version oder jede spätere Version dafür gilt, haben Sie die Wahl den Bestimmungen dieser speziell benannten Version zu folgen, oder jeder Version, die später von der Free Software Foundation, nicht als Entwurf, veröffentlicht wurde.