

# Teil 13: Einführung in das Internet

## Literatur:

- Rainer Klute: Das World Wide Web. Addison-Wesley, 1996, ISBN: 389319763X.
- RRZN Hannover: Internet. Ein Einführung in die Nutzung der Internet-Dienste. Es gibt inzwischen die 7. Auflage. Erhältlich bei Beratung des HRZ.
- Douglas Comer: Internetworking with TCP/IP. Prentice Hall, 1988, ISBN 0134701887.
- W. Richard Stevens: TCP Illustrated, Vol. 1. Addison-Wesley, 1994, ISBN 0201633469.
- Craig Zacker: Upgrading and Troubleshooting Networks — The Complete Reference. Osborne/McGraw-Hill, 2000, ISBN 0-07-212256-0, 918 pages.
- Jochen Musch: Die Geschichte des Netzes: ein historischer Abriß.  
<http://www.psychologie.uni-bonn.de/sozial/staff/musch/history.htm>
- Gregory R. Gromov: The Roads and Crossroads of Internet History.  
<http://www.internetvalley.com/intval1.html>
- Robert H'obbes' Zakon: Hobbes' Internet Timeline v5.1  
<http://info.isoc.org/guest/zakon/Internet/History/HIT.html>
- Tim Berners-Lee: Weaving the Web. Harper, 1999, ISBN: 0062515861, 226 pages.
- Dan Connolly, Robert Cailliau: A Little History of the World Wide Web.  
<http://www.w3.org/History.html>

# Lernziele

Nach diesem Kapitel sollten Sie Folgendes können:

- erklären, warum man das Internet auch “ein Netzwerk von Netzwerken” nennt.
- den Aufbau numerischer IP-Adressen erklären.
- den Begriff “Protocol Stack” erklären.
- die Bedeutung von Port Nummern erklären.
- die Funktionsweise des “Domain Name System” erklären.
- sich Informationen über eine IP-Adresse oder ein Domain Namen verschaffen.

# Inhalt

1. Das Internet

2. Geschichte von Internet und WWW

3. Protokoll Schichten ( "Protocol Stack" )

4. Domain Name System

5. Internet Hosting (Eigene Domain)

# Das Internet (1)

- Das Internet ist ein System von mit einander verknüpften Netzen, die die Protokolle der TCP/IP Familie nutzen.

Ein Protokoll definiert die Syntax für den Nachrichtenaustausch zwischen Computern. IP: "Internet Protocol". TCP: "Transmission Control Protocol".

- Im Januar 2009 waren ca. 625 Millionen Rechner an das Internet angeschlossen

Quelle: [<https://www.isc.org/solutions/survey/>].

Niemand weiß es genau. Diese Zahl ist die Anzahl Abbildungen von IP-Nummern auf Namen im Domain Name System.

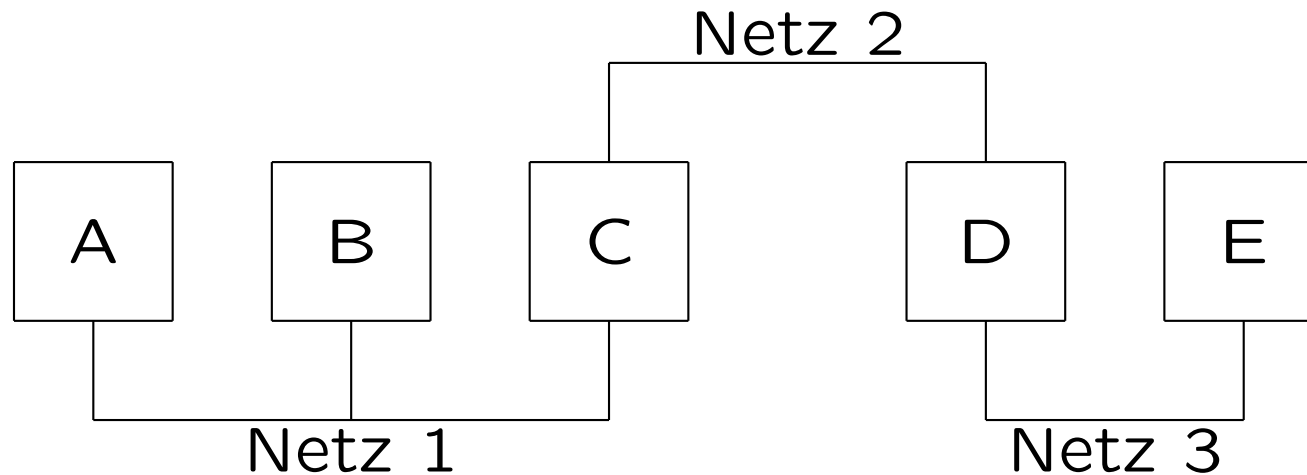
Jan 2000: 72 million, Juli 1995: 5–8 million, October 1990: ~313000.

.com: 95 mio, .net: 190 mio, .edu: 11 mio, .org: 1 mio, .de: 21 mio.

## Das Internet (2)

- Obwohl das Wort “Internet” heute oft synonym zum “World Wide Web” verwendet wird, ist das WWW nur eine Anwendung des Internets.
- Das Internet bietet auch andere Dienste: Email, FTP, Telnet, News, Finger, Talk, DNS, NFS, ...
- Web Browser unterstützen einige dieser Dienste, und ein wichtiger Faktor für den Erfolg des WWW war seine integrierende Funktion.
- Web Server werden über das Protokoll HTTP angesprochen.

## “Netz von Netzen”



- Wenn Rechner A Daten an Rechner E schicken will, müssen die Datenpakete von Rechner C und D weitervermittelt werden.

Abhängig von der Protokollebene, auf der dies geschieht, heißen Rechner C and D: “Bridge”, “Router”, oder “Gateway”.

# Internet Adressen (1)

- Internet Adressen (v4) sind 32-Bit Zahlen, geschrieben als Folge von vier Bytes, z.B. 141.48.14.50.

Jedes Byte kann die Werte 0–255 annehmen. Inzwischen nimmt die Verbreitung von IPv6 zu, dort sind die Internet Adressen 128 Bit lang.

- Sie werden auch IP Nummern genannt, weil sie im Internet Protokoll (IP) zur Identifikation von Rechnern benutzt werden.
- Namen wie z.B. `haendel.informatik.uni-halle.de` werden vom “Domain Name System” (s.u.) in die intern verwendeten Nummern übersetzt.

## Internet Adressen (2)

- Es ist nicht möglich, daß zwei Rechner im Internet gleichzeitig die gleiche IP Nummer haben.

Wenn Rechner nur manchmal mit dem Internet verbunden sind (man sich z.B. über Modem einwählt), kann ihnen dynamisch eine IP Nummer zugewiesen werden. Nachdem Ende der Sitzung kann die gleiche Nummer einem anderen Rechner zugewiesen werden.

- Ein Rechner kann mehrere IP Nummern haben.

Z.B. wenn er mehrere Netzwerkkarten hat. Genauer identifiziert die IP Nummer nicht den Rechner, sondern den Netzwerkanschluß.

- Eine spezielle IP Nummer ist 127.0.0.1 (`localhost`): Dies ist immer der eigene Rechner.

## Internet Adressen (3)

- Internet Adressen bestehen aus einem Netzwerk-Anteil und einem Rechner-Anteil (“host”, identifiziert den Rechner innerhalb des lokalen Netzes).

Früher wurden verschiedene Klassen von IP Adressen unterschieden: Klasse A (erstes Byte 0–127): 1 Byte Netzwerk / 3 Byte Rechner (für große Netze). Klasse B (erstes Byte 128–191): 2 Byte/2 Byte. Klasse C (erstes Byte 192–223): 3 Byte/1 Byte. Da die Internet Adressen knapp geworden sind, wurde dieses Schema aufgegeben, und die Grenze kann nun bei jeder Bitposition liegen. Man schreibt z.B. 141.48.0.0/14 (14 Bit Netzwerk, 18 Bit Rechner im Netz).

- Oder: Netzwerk, Subnetz, Host (s.u.).

Aus globaler Sicht gehört das Subnetz zum Host-Anteil, aus lokaler Sicht zum Netzwerk-Anteil. Siehe Folie 13-12.

## Internet Adressen (4)

- Für das globale Routing im Internet (Verteilung von Datenpaketen an Verzweigungspunkten) ist nur der Netzwerk-Anteil wichtig.
- Wären die IP-Nummern völlig beliebig verteilt, so wären die Routing-Tabellen viel größer.
- Wenn ein Rechner an ein anderes Netzwerk angeschlossen wird, braucht er eine neue IP-Nummer.

## Internet Adressen (5)

- Die Universität Halle hat vom DFN Verein (Deutsches Forschungsnetz) folgende Netzwerk-Adresse zugewiesen bekommen: 141.48.0.0/16.
- Man bekommt seine IP Nummer(n) von seinem ISP (“Internet Service Provider”: Firmen, von denen man einen Internet-Anschluß bekommen kann).

Offiziell werden IP Nummern von der IANA vergeben (“Internet Assigned Numbers Authority”) [<http://www.iana.org>]. Sie hat die Arbeit auf drei regionale Agenturen verteilt: APNIC für Asien/Pazifik, ARIN für Amerika, und RIPE NCC für Europa [<http://www.ripe.net>]. Das RIPE NCC verteilt die Aufgaben wiederum auf “Local Internet Registries” (LIRs). Der DFN Verein [<http://www.dfn.de>] ist wie alle großen ISPs eine LIR.

## Internet Adressen (6)

- Alle IP-Nummern der Universität Halle beginnen mit 141.48. Die hinteren 16 Bit werden vom Rechenzentrum vergeben.
- Da das Gesamtnetz der Universität wieder aus vielen kleinen Netzen besteht (z.B. eins pro Institut), ist dieser Bereich noch in einen Subnetz-Anteil und den eigentlichen Rechner-Anteil aufgeteilt.

Z.B. 10 Bit zur Subnetz-Identifikation und 6 Bit für den Rechner im Subnetz, aber das variiert je nach Institutsgröße. Früher gab es für jedes Gebäude ein Subnetz. Heute werden mit "Switches" VLANs ("virtual local area networks") gebildet. Z.B. sind alle Rechner der Bibliothek in einem Subnetz, egal in welchem Gebäude sie stehen.

## Internet Adressen (7)

- Wenn ein Rechner  $A$  mit einem Rechner  $B$  kommunizieren möchte, ist wichtig zu wissen, ob der Zielrechner ( $B$ ) im gleichen lokalen Netz ist.

Dann kann  $A$  ihn direkt ansprechen. Ansonsten muß er die Nachricht nicht direkt an  $B$  schicken, sondern an einen Router (der einen Anschluß im lokalen Netz hat und den Verkehr nach außen weiterleitet).

- Dazu müssen Netz- und Subnetz-Anteil der Adressen von  $A$  und  $B$  übereinstimmen.

Bei der Netzwerk-Konfiguration muß oft eine Subnetz-Maske angegeben werden. Die hat 1-Bits für Netz- und Subnetz-Anteil, da zwischen diesen Komponenten lokal nicht unterschieden werden muß. Bei 16 Bit Netzwerk, 10 Bit Subnetz, und 6 Bit Rechner-Anteil ist die Subnetz-Maske also 255.255.255.192 (das letzte Byte ist 11000000).

# Anschluß der Universität

- Die Universität ist mit 1 GBit/s an das X-WiN angeschlossen, das vom DFN Verein verwaltet wird.

Die Verbindung ist zum nächsten Kernnetzknotten in Leipzig. Halle hat einen Cluster-Anschluß zusammen mit der Hochschule Merseburg, d.h. es gibt nur einen gemeinsamen Anschlußpunkt an das X-WiN. Die maximale Bandbreite von 1 GBit/s bezieht sich nur auf den Anschluß an das X-WiN. Wenn es viel Netzwerkverkehr im X-WiN oder vom X-WiN in das globale Internet gibt, ist die tatsächliche Bandbreite geringer.

- Die Universität bezahlt 203.847 Euro pro Jahr an den “DFN Verein” (Anschlußtyp I10, cluster).

Das sind 558 Euro pro Tag (auch am Wochenende). Der Minutenpreis ist 39 Cent.

# X-WiN (1)

- Der DFN Verein hat mehr als 300 Mitglieder (Universitäten, Forschungseinrichtungen).
- Es gibt zweimal im Jahr eine Betriebstagung, und die Folien der Vorträge sind online.

[<http://www.dfn.de/veranstaltungen0/betriebstagung/vortraege-bt/>].

Eine aktuelle Karte des X-WiN findet sich in

[<http://www.dfn.de/fileadmin/3Beratung/Betriebstagungen/bt48/plenum-xwin-hma.pdf>]

Siehe auch: [<http://www.aufdemhoevel.de/xwin.htm>]

- Kernnetz: Multi Gigabit, ungefähr 50 Knoten.

Im Januar 2008 war das übertragene Datenvolumen 4641 TB (Daten, die in diesem Monat aus dem Netz “exportiert” wurden).

## X-WiN (2)

- Für viele Verbindungen hat der DFN Verein die nackte Glasfaser (“dark fiber”) gemietet, und betreibt seine eigene Hardware, um Daten durch die Faser zu schicken.

Die Bandbreite von Fasern können mit neueren WDMs (“wavelength division multiplexer”) gesteigert werden. Z.B. werden 16 or 32 verschiedene Wellenlängen (Lichtfarben) genutzt, jede mit 10 GBit/s. Die Glasfaser ist u.a. von KPN und von GasLINE gemietet. Weil Gasleitungen meist gut vor Schäden durch Bauarbeiten geschützt sind, ist dies auch ein sicherer Platz für Kommunikationsleitungen. Ein typisches Glasfaserkabel enthält 144 Fasern. Für einige Verbindungen hat der DFN Verein auch nur eine Wellenlänge in einer Glasfaser gemietet, die von einer anderen Firma betrieben wird. Kernnetz-Knoten sind mit mindestens zwei anderen Kernnetz-Knoten verbunden.

## X-WiN (3)

- Verbindungen des DFN zu anderen Netzen (erst dadurch wird es Bestandteil des Internets):
  - ◇ via GÉANT (DANTE) zu europäischen und nord-amerikanischen Forschungsnetzen,
  - ◇ via DE-CIX zu kommerziellen Internet Providern in Deutschland,
    - DE-CIX ist in Frankfurt [<http://www.decix.de>]. Es ist ein “Peering Point”, d.h. die Netzbetreiber haben ihre Netze zum gegenseitigen Vorteil und ohne Bezahlung verknüpft. X-WiN hat noch zusätzliche Verbindungen zu T-Internet und lokalen ISPs.
  - ◇ via Global Crossing und TELIA weltweit (“global upstream”).

# Traceroute (1)

- `tracert` (unter Windows) zeigt die Rechner an, die die Daten auf dem Weg vom eigenen Rechner zum Zielrechner weiterleiten (Router).

Man kann dieses Kommando nicht nur vom eigenen Rechner aus verwenden: [<http://www.traceroute.org>].

- Beispiel: `tracert www.tu-clausthal.de` zeigt eine Verbindung von der MLU Halle zur TU Clausthal über Leipzig, Essen, Hannover, Göttingen.

Zumindest, als ich die Ausgabe (siehe nächste Folie) noch bekommen konnte (2004). Inzwischen sperrt eine Firewall diesen Dienst (zum Schutz vor Hacker-Angriffen werden bestimmte Arten von Nachrichten nicht mehr zwischen Universitätsnetz und DFN weitergeleitet).

## Traceroute (2)

```
traceroute www.tu-clausthal.de
auf haendel.informatik.uni-halle.de (141.48.14.50):
```

```
 1 141.48.14.254          (141.48.14.254) 0.3 ms
 2 141.48.0.254          (141.48.0.254)  0.4 ms
 3 wincisco.urz.uni-halle.de (141.48.25.10)  0.4 ms
 4 ar-leipzig2.g-win.dfn.de  (188.1.35.5)   3.4 ms
 5 cr-leipzig1-ge8-0.g-win.dfn.de (188.1.70.1)   3.1 ms
 6 cr-essen1-po1-0.g-win.dfn.de (188.1.18.105) 12 ms
 7 cr-hannover1-po0-2.g-win.dfn.de (188.1.18.50)  17 ms
 8 ar-goettingen1-po0-0.g-win.dfn.de (188.1.88.74)  19 ms
 9 7200v.xr.rz.tu-clausthal.de (139.174.251.9) 21 ms
10 sr-backbone.rz.tu-clausthal.de (139.174.254.2) 24 ms
11 *
12 *
```

## Traceroute (3)

- Es gibt allerdings nicht “die” Verbindung. Wenn ein Stau oder eine Störung auftritt, können Datenpakete sofort auch einen anderen Weg nehmen.

Außerdem werden die Routingtabellen ständig aktualisiert und neuen Gegebenheiten angepasst.

- Traceroute druckt “\*”, wenn es (in 5 Sek.) keine Antwort von einer bestimmten Netzwerk-Distanz (Anzahl Weiterleitungen) bekommen hat.

Sonst zeigt es die Zeit bis zum Eintreffen der Antwort an, und probiert jede Entfernung drei Mal (auf der Folie ist nur eine Zeit gezeigt). Wenn es kurzzeitig einen “Stau auf der Datenautobahn” gab, können die Zeiten bei längeren Entfernungen auch mal kürzer sein.

# Traceroute (4)

traceroute paradox.sis.pitt.edu aus Gießen im Jahre 2000:

1	cisgis.uni-giessen.de	(134.176.24.40)	0.3 ms
2	cisgis-vlan-1.uni-giessen.de	(134.176.253.1)	1.4 ms
3	ar-marburg2.g-win.dfn.de	(188.1.42.133)	2.1 ms
4	ar-marburg1.g-win.dfn.de	(188.1.81.1)	1.9 ms
5	cr-frankfurt1.g-win.dfn.de	(188.1.80.45)	3.4 ms
6	ir-frankfurt2.g-win.dfn.de	(188.1.80.38)	3.4 ms
7	dfn.de1.de.geant.net	(62.40.103.33)	4.9 ms
8	de1-1.de2.de.geant.net	(62.40.96.130)	3.7 ms
9	62.40.103.254	(62.40.103.254)	84.6 ms
10	clev-nycm.abilene.ucaid.edu	(198.32.8.29)	96.8 ms
11	abilene.psc.net	(192.88.115.122)	100.1 ms
12	pitt.psc.net	(198.32.224.8)	102.7 ms
13	136.142.2.101	(136.142.2.101)	103.2 ms
14	paradox.sis.pitt.edu	(136.142.116.28)	104.8 ms

# Weitere Kommandos (1)

- `ping` prüft, ob ein Rechner erreichbar ist, z.B.

```
ping www.acm.org  
→ www.acm.org is alive
```

Wenn eine Netzwerkverbindung nicht klappt, kann man zuerst mit “`ping localhost`” prüfen, ob auf dem eigenen Rechner die Netzwerk-Software richtig läuft. Dann kann man schrittweise weiter entferntere Rechner prüfen.

Unter Windows schickt `ping` mehrere Test-Datenpakete und druckt jeweils die Laufzeit. Unter UNIX druckt es nur “is alive”. Man kann dort aber mit “`ping -s www.acm.org`” einen Test pro Sekunde machen, bis `Ctrl+C` gedrückt wird. Dann bekommt man auch eine Statistik. Es gibt noch viele weitere Optionen, siehe “`ping`” (ohne Optionen) unter Windows und `man ping` unter UNIX. Z.B. zeigt `ping -r 9 www.dfn.de` unter Windows die Route (bis 9 Stationen) an.

## Weitere Kommandos (2)

- `netstat -r`

zeigt die Routing Tabelle des lokalen Rechners.

Unter Windows können Sie alternativ auch `route PRINT` verwenden.

- Z.B. liefert `netstat -rn` auf `141.48.14.50`:

Destination	Gateway	Flags	Ref	Use	Interface
141.48.14.0	141.48.14.50	U	1	7380	eri0
224.0.0.0	141.48.14.50	U	1	0	eri0
default	141.48.14.254	UG	1	16025	
127.0.0.1	127.0.0.1	UH	5	527	lo0

Die Option `-n` bedeutet, daß numerische IP-Adressen statt Rechnernamen benutzt werden. Natürlich ist die Ausgabe interessanter auf einer Maschine, die mehrere Netzwerk-Anschlüsse hat. `141.48.14.254` ist der Router, der das lokale Netz mit dem Universitätsnetz verknüpft (und von dort aus weiter). `224.0.0.0` ist eine "multicast" Adresse (`BASE-ADDRESS.MCAST.NET`). Unter "Flags" bedeutet U: "up", G: "gateway", H: "local host" (?).

## Weitere Kommandos (3)

- `netstat` zeigt alle Netzwerkverbindungen an, die im Moment auf dem Rechner bestehen.

Unter Windows versuchen Sie auch `netstat -a` (alle).

- Informationen über Netzwerk-Schnittstellen erhalten Sie mit:
  - ◇ `netstat -i` (UNIX)
  - ◇ `ifconfig -a` (UNIX)
  - ◇ `ipconfig` oder `ipconfig /All` (Windows)

# Geschwindigkeitsmessungen

Allgemeiner Zustand des Internets:

- [<http://www.internettrafficreport.com/>]
- [<http://weather.ucr.edu/>]

“Internet Weather Report” .

Messungen von der University of California at Riverside.

Andere Verbindungen: [<http://www.noc.ucla.edu/weather.html>]

Geschwindigkeit Ihrer Internet-Verbindung:

- [<http://de.netmeter.eu/>]
- [<http://www.speedmeter.de/speedtest>]

# Inhalt

1. Das Internet

2. Geschichte von Internet und WWW

3. Protokoll Schichten ( "Protocol Stack" )

4. Domain Name System

5. Internet Hosting (Eigene Domain)

# Geschichte des Internet (1)

- 1958: ARPA in den USA gegründet.

ARPA or DARPA: “Advanced Research Projects Agency” des “US Department of Defense”. Reaktion auf den Sputnik-Schock 1957.

- 1968: Vertrag für ARPANET Hardware mit BBN.

IMPs: Interface Message Processors.

[<http://www.zakon.org/robert/internet/timeline/>].

- 1969: ARPANET startet mit 4 Universitäten.

University of California at Los Angeles (UCLA), University of California at Santa Barbara (UCSB), Stanford Research Institute (SRI), University of Utah. Jede hatte eine andere Hardware. Die Verbindung war über die IMPs und Telefonleitungen von AT&T mit 50 KBit/s.

## Geschichte des Internet (2)

- 1972: ARPANET mit ungefähr 40 Computern wurde auf der “First International Conference on Computer Communications” vorgestellt.

Dienste des ARPANET: Remote Login, File Transfer, Electronic Mail. EMail wurde 1971/1972 eingeführt (nach telnet and ftp). Nach einer Studie von 1973 war Email 75% des ARPANET Verkehrs.

- 1972: InterNetwork Working Group gegründet.

Ziel: Gemeinsames Protokoll für internationale Verbindungen zwischen autonomen Netzen. Vinton Cerf war erster Vorsitzender.

- 1973/74: TCP Protokoll spezifiziert (Kahn/Cerf).

Die Aufteilung in TCP und IP geschah später (1978).

# Geschichte des Internet (3)

- 1976/77: UUCP: Unix-to-Unix-Copy Protocol.

Entwickelt von AT&T Bell Labs (Mike Lesk and others), enthalten in UNIX V7. "Poor Man's ARPANET".

- 1979: USENET/News (basierend auf UUCP).

- 1981: CSNET gegründet.

Nur Universitäten mit ARPA Projekten hatten ARPANET-Zugang. CSNET ("Computer Science Network", "Computer and Science Network") bot auch anderen Wissenschaftlern Netzwerk-Dienste.

- 1981: BITNET gegründet.

Basierend auf IBM's Network Job Entry protocol.

- 1982: EUnet (European UNIX Network) gegründet.

# Geschichte des Internet (4)

- 01.01.1983: ARPANET wechselt auf TCP/IP.
- 1983: Gateway zwischen CSNET und ARPANET.
- 1983: EARN (European Academic and Research Network): Ableger des BITNET (mit Gateway).
- 1983: MILNET vom ARPANET getrennt (68/113)
- 1983: Quellcode der TCP/IP Implementierung in 4.2 BSD wird öffentlich verfügbar gemacht.

Entwickelt von der University of California at Berkeley mit Geld von der DARPA. Viele Betriebssystem-Hersteller verwenden diese Implementierung (fördert Verbreitung und Portabilität).

# Geschichte des Internet (5)

- 1983/84: Name Server, Domain Name System
- 1986: NSFNET gegründet (Backbone 56 Kbit/s).
- 1988: EUnet wechselt auf TCP/IP.
- 1988: NSFNET Backbone T1 (1.544 Mbit/s).
- 1989: Mehr als 100 000 Rechner sind mit dem Internet verbunden.
- 1990: ARPANET wird aufgelöst. NSFNET übernimmt seine Rolle als Herzstück des Internet.

# Geschichte des WWW (1)

- 1945: Vannevar Bush schlägt eine Maschine für Verweise zwischen Dokumenten auf Microfilm vor.
- 1963–68: Doug Engelbart entwickelt das “oNLine System” (NLS), und erfindet die Maus dafür.
- 1965: Ted Nelson prägt den Begriff “Hypertext”.
- 1979: Charles Goldfarb entwickelt SGML.
- 1980: Tim Berners-Lee entwickelt am CERN ein Hypertext Programm “Enquire”.

CERN: Europäisches Zentrum für Kernphysik/Teilchenphysik in Genf.  
Er machte dort ein Praktikum.

## Geschichte des WWW (2)

- 1987: CERN wird ans Internet angeschlossen.
- 1989: Tim Berners-Lee schreibt einen Antrag für ein globales Hypertext Projekt am CERN.

Robert Caillian schrieb unabhängig davon einen ähnlichen Antrag.  
1990 entwickelten Sie gemeinsam eine neue Version.

- 1990: Von Oktober bis Dezember entwickelte Tim Berners-Lee den ersten Web-Server "httpd" und den ersten Browser "WorldWideWeb".
- 1991: Das Programm "WorldWideWeb" wurde im Internet öffentlich verfügbar gemacht.

# Geschichte des WWW (3)

- 1991–93: Tim Berners-Lee koordiniert die weitere Entwicklung von URIs, HTTP, und HTML.

Im Dezember 1991 wird das System auf einer Konferenz vorgestellt. Im Mai 1992 erscheint ein Artikel von Tim Berners-Lee in den “Proceedings of the Third Joint European Networking Conference”, der URLs, HTTP, und HTML beschreibt. Die HTML Version von damals war noch recht einfach. Sie enthielt nur Überschriften verschiedener Ebenen, Listen (geordnet, ungeordnet, Definitionslisten), Hypertext-Links, Titel, Adresse (und eventuell vorformatierten Text, IsIndex). Dave Raggett schlägt Ende 1991 erweiterte Version HTML+ vor.

- 1993 (Januar): Ungefähr 50 WWW/HTTP Server.

Im Oktober sind es schon über 200, im Juni 1994 über 1500, im November 1995 ungefähr 73500.

# Geschichte des WWW (4)

- 1993: Web-Browser Mosaic veröffentlicht.

Mosaic bot “eine konsistente und einfach zu nutzende Hypertext-basierte Oberfläche zum Zugriff auf sehr unterschiedliche Informationsquellen, u.a. Gopher, WAIS, World Wide Web, NNTP/News, Techinfo, Textinfo, FTP, lokale Dateisysteme, telnet, tn3270.” Mosaic wurde von Marc Andreessen and Eric Bina am NCSA entwickelt (National Center for Supercomputing Applications in den USA). Mosaic war für X, PCs, und Macintosh verfügbar. Wegen der öffentlichen Gelder für das NCSA wurde der Quellcode Public Domain.

- 1993: NCSA bietet auch einen Web Server, HTTPD.

Im Dezember 1993 erschien ein langer Artikel über das WWW und Mosaic in der The New York Times. Am Jahresende wurden mehr als 1000 Kopien von Mosaic und NCSA HTTPD pro Tag vom NCSA heruntergeladen.

# Geschichte des WWW (5)

- 1994 (Mai): Erste internationale WWW Konferenz.

Die Konferenz wurde am CERN abgehalten. Schon im Oktober des gleichen Jahres fand die zweite Konferenz statt (in Chicago).

- 1994: Marc Andreessen und andere verlassen das NCSA und gründen die “Mosaic Communications Corp” (später Netscape).

- 1994: Das W3 Consortium wird gegründet, mit Tim Berners-Lee als Direktor [<http://www.w3.org>].

Das W3C ist am MIT angesiedelt. Das CERN hatte sich gerade für einen neuen Beschleunigerring entschieden und kein Geld, das WWW weiter zu entwickeln. Das W3C ist über Mitgliedsbeiträge finanziert.

# Geschichte des WWW (6)

- 1995: Sun Microsystems stellt HotJava vor, einen Browser, der interaktive Objekte enthält.
- 1995: Microsoft kauft einen Browser, der von der Firma Spyglass entwickelt wurde (und Mosaic Code enthält), und macht daraus den Internet Explorer.
- 1995 (Nov.): HTML 2.0 wird als “Internet Proposed Standard” verabschiedet (RFC 1866).

Dies war die erste Version, die formal über eine SGML DTD definiert war. RFCs (“Request for Comments”) sind Internet Standards (und andere Internet Dokumente) siehe z.B. [<http://www.rfc-editor.org>].

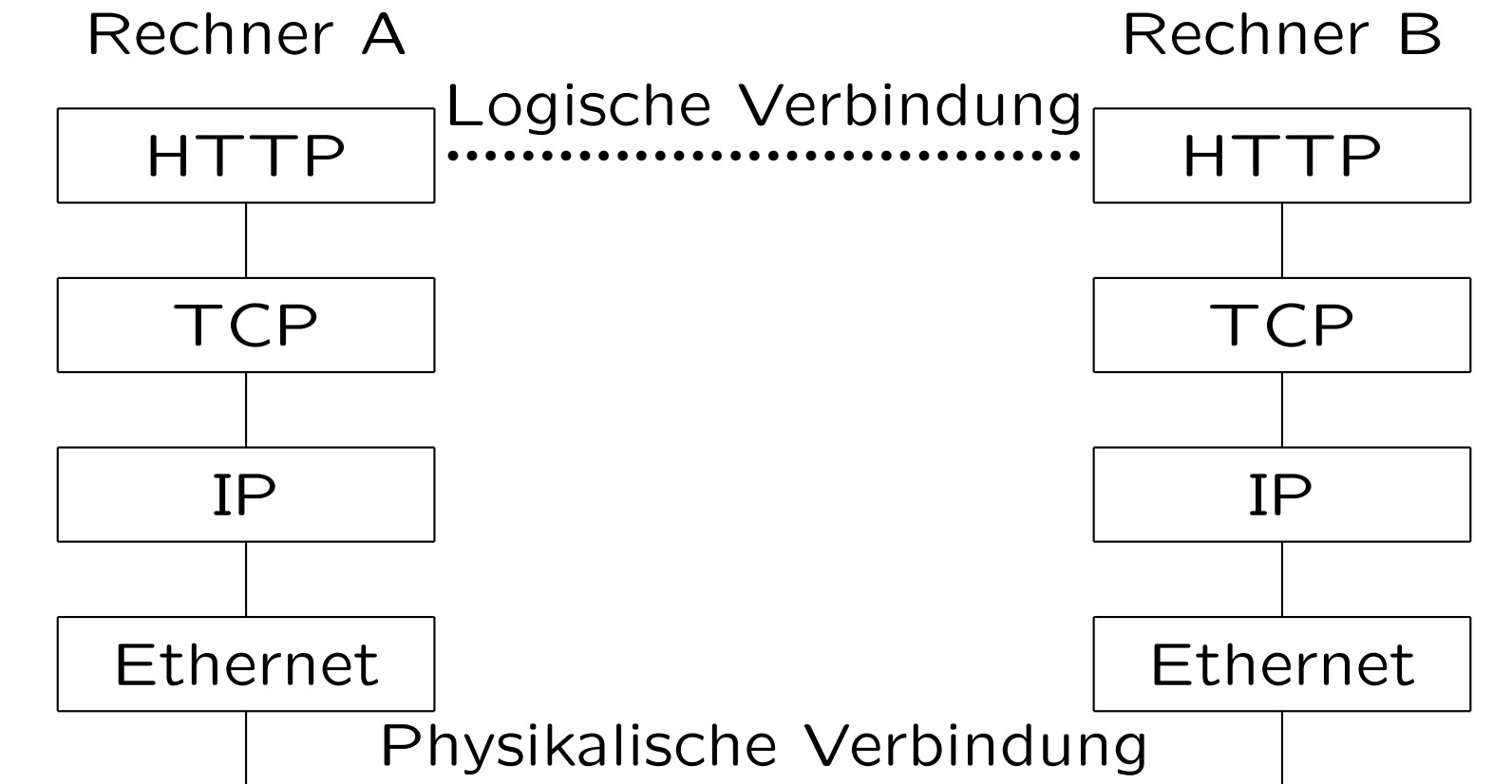
# Inhalt

1. Das Internet
2. Geschichte von Internet und WWW
3. Protokoll Schichten ( "Protocol Stack" )
4. Domain Name System
5. Internet Hosting (Eigene Domain)

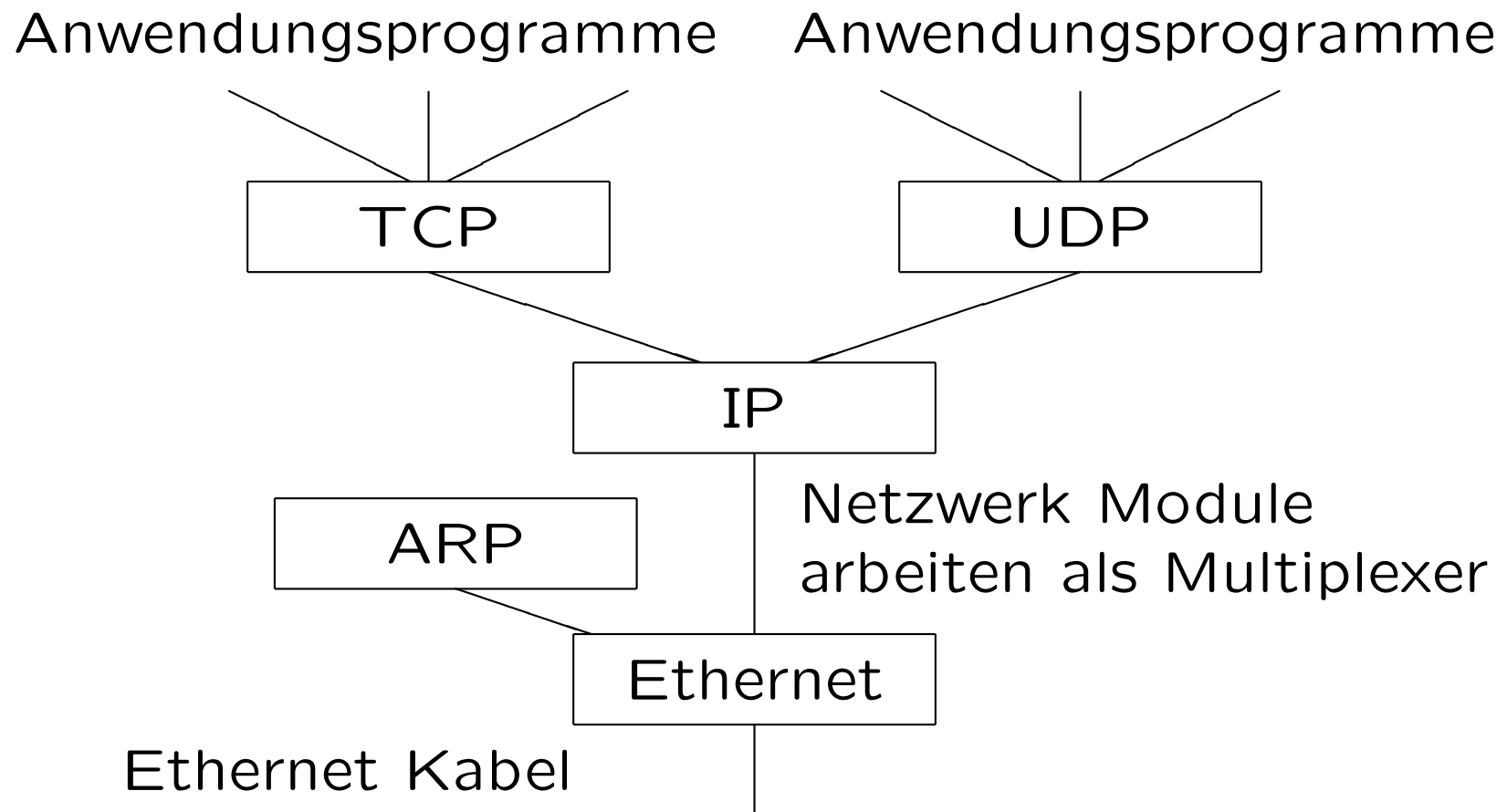
# Schichtenarchitektur

- Es ist ein allgemeines Prinzip in der Informatik, komplexe Systeme schichtenweise aufzubauen.
- Jede Schicht benutzt die Dienste der darunterliegenden Schicht, um selbst mächtigere Dienste zur Verfügung zu stellen.
- Die Schnittstellen der Schichten abstrahieren dabei immer weiter von bestimmten Problemen (die in der jeweiligen Schicht behandelt werden).
- Man kann Schichten bei Bedarf austauschen.

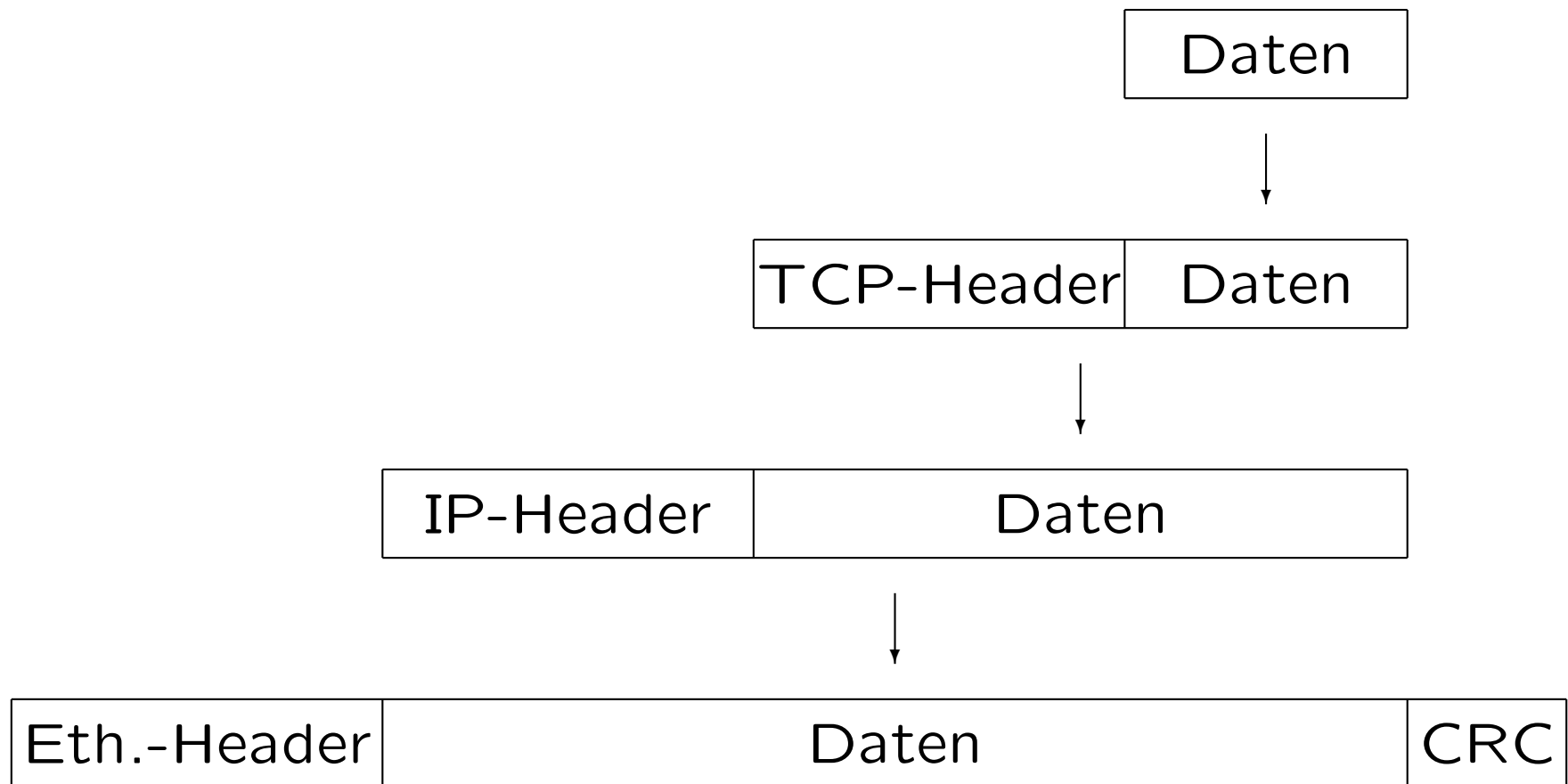
# “Protocol Stack”



# Protokolle auf einem Rechner

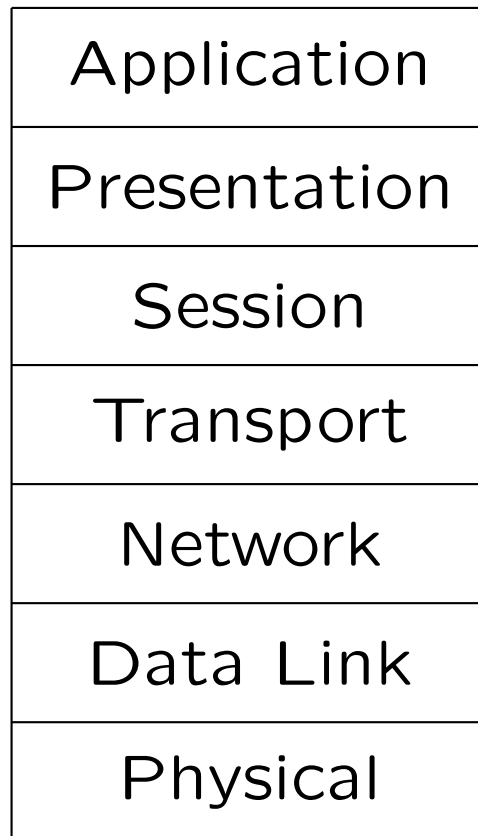


# Verpacken von Datenpaketen

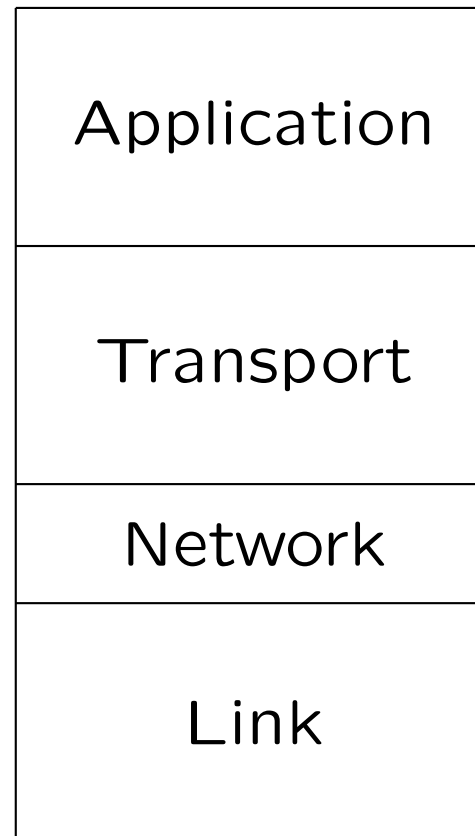


# Vergleich mit OSI-Model

## OSI



## TCP/IP



Telnet, FTP,  
HTTP, etc.

TCP, UDP

IP, ICMP

Device Driver  
Hardware

# Ethernet (1)

- Zum Beispiel ist Ethernet ein typisches Protokoll für lokale Netze (innerhalb eines Gebäudes).
- Ethernet kann Datenpakete (bis 1500 Byte) zwischen zwei Rechnern übertragen, die (mehr oder weniger) direkt miteinander verbunden sind.

Früher ging ein Koaxialkabel an jedem Rechner im Netz vorbei. Ethernet hat eine Bus-Topologie. D.h. ein Rechner sendet auf dem Kabel, die anderen horchen. Normalerweise werden alle Datenpakete ignoriert, die nicht für den eigenen Rechner bestimmt sind, aber Hacker können alle Datenpakete mitlesen. Deswegen ist es so wichtig, daß Daten, insbesondere Passwörter, verschlüsselt übertragen werden. Heute wird Ethernet meist über ein "Twisted Pair" Kabel (Telefonkabel) über einen zentralen "Hub" oder "Switch" verbunden.

## Ethernet (2)

- Computer sind im Ethernet über 48 bit Adressen identifiziert, z.B. 8:0:20:12:5a:f6.

Ethernet Adressen werden üblicherweise als 6 Bytes hexadezimal geschrieben (d.h. zur Basis 16), durch ":" getrennt. Allgemeiner werden global eindeutige Adressen in Netzwerkprotokollen auf der OSI Schicht 2 (dazu gehört das Ethernet) MAC (Media Access Code) oder MAC (MA Control) Adressen genannt. Man kann Ethernet-Adressen mit dem Kommando `arp` herausfinden.

- Die Adressen sind in der Ethernetkarte festgelegt.

Die ersten drei Bytes bekommt der Hersteller von der IEEE zugewiesen, die anderen drei vergibt er selbst. Bei vielen Karten ist die Adresse aber konfigurierbar (interessant für Hacker). Wenn ein Computer eine neue Netzwerk-Karte bekommt, bekommt er automatisch eine neue Ethernet-Adresse. Die IP Adresse kann er aber behalten.

## Ethernet (3)

- Ethernet war ursprünglich mit 10 Mbit/s definiert.

Gelegentlich noch in alten Gebäuden verwendet, Koaxialkabel wird von neueren Versionen nicht unterstützt. Übrigens ist es wichtig, Kabelenden mit 50  $\Omega$ -Widerständen zu terminieren. Sonst kommt es dort zu Reflektionen, die das Signal zerstören.

- Üblich: 100 Mbit/s ( "Fast Ethernet", 100BaseTX).

Es kam ca. 1996 auf den Markt und benutzt üblicherweise ein Twisted Pair Kabel (Category 5 UTP, RJ-45 Stecker). Zwei Computer kann man mit einem überkreuzten Kabel verbinden, sonst braucht man einen Hub oder Switch. 2002 konnte ein PC via FTP maximal 30–40 Mbit/s übertragen, das Netzwerk war dann nicht der Flaschenhals. Ein Film in Fernsehqualität braucht 3–4 Mbit/s. [c't 3/2002].

- Inzwischen: Gigabit Ethernet (und sogar 10 GB/s).

# IP: Internet Protokoll (1)

- Das Internet-Protokoll IP befördert Datenpakete (Folgen von Bytes, maximale Länge 64 KB) von einem Quellrechner zu einem Zielrechner.
- Quelle und Ziel sind durch 32 Bit IP-Adressen identifiziert (s.o.).
- Die Zustellung ist nicht garantiert: Datenpakete können verloren gehen.

Es gibt nicht einmal sicher eine Information über den Fehler. Natürlich werden Datenpakete nicht völlig beliebig weggeworfen, sondern nur in Fehlerfällen ("best effort"). Das darüberliegende Protokoll TCP macht die Zustellung sicher (s.u.).

# IP: Internet Protokoll (2)

- Hauptaufgabe von IP: Weiterleitung zwischen verschiedenen Netzen, Routing auf diesem Weg.
- Routing Entscheidungen werden lokal getroffen.

Es gibt keine Instanz im Netz, die den gesamten Weg von der Quelle zum Ziel plant. Jeder Router wählt nur die nächste Station.
- Verbindungslos: Jedes Paket wird einzeln behandelt.

Es ist möglich, daß Pakete einer längeren Nachricht verschiedene Wege durch das Netz nehmen.
- Die IP-Ebene fragmentiert ggf. Datenpakete, die zu groß für die Schicht darunter sind.

Und setzt die Fragmente am Ziel wieder zusammen.

# IPv6 (1)

- Die heute verbreitete Version von IP ist Version 4.

Version 5 war kein Nachfolger, sondern ein experimentelles Streaming-Protokoll für Audio und Video.

- Die Beschränkung auf 32 Bit Adressen (ca. 4 Milliarden) ist ein wesentliches Problem.

Vieles wird einfacher und effizienter, wenn man einen großen Adressraum hat, in dem man Teile für bestimmte Aufgaben oder zukünftige Rechner reservieren kann, und nicht gezwungen ist, und mehr oder weniger jede verfügbare Adresse wirklich zu nutzen.

- In IPv6 (IP Version 6, auch “IP next generation: IPng” genannt) sind die Adressen 128 Bit lang.

## IPv6 (2)

- IPv6 wurde 1994 von der Internet Engineering Task Force als “Proposed Standard” verabschiedet.

1998 wurde es “Draft Standard”. [<http://www.ietf.org/>]

- Es gibt einige IPv6 Netze, und mehr und mehr Hard- und Software unterstützt IPv6.

Software muß an die längeren Adressen angepasst werden. Windows XP unterstützt auch IPv6.

- Für die absehbare Zukunft werden beide Protokolle nebeneinander benutzt werden.

# TCP (1)

- TCP: “Transmission Control Protocol” .
- Eine TCP-Verbindung ist wie telefonieren:
  - ◇ Ein Client baut eine Verbindung zu einem Server auf (wählt eine Nummer).
  - ◇ Dann gibt es Datenströme in beiden Richtungen.

Es handelt sich um eine “full duplex” Verbindung, d.h. gibt zwei unabhängige Datenströme (Folgen von Bytes) von A nach B und von B nach A. Paketgrenzen sind nicht mehr sichtbar. Natürlich können beide Seiten auch “zuhören”, d.h. auf bestimmte Daten von der anderen Seite warten, bevor sie wieder Daten senden.
  - ◇ Jede von beiden Seiten kann die Verbindung beenden ( “auflegen” ).

## TCP (2)

- Während IP Datenpakete nur an einen Rechner schickt, erlaubt TCP Verbindungen zu einem bestimmten Programm auf dem Rechner.
- Die “virtuellen Netzwerkanschlüsse” eines Rechners werden über 16 Bit Portnummern unterschieden.
- Wenn ein Serverprozess startet, teilt er dem Betriebssystem mit, auf welchem Port er auf “Anrufe” warten will.

Natürlich muß der Port noch frei sein. Ein Webserver benutzt typischerweise Port 80. Das ist aber konfigurierbar, z.B. wird teils auch 8080 genutzt.

## TCP (3)

- Wenn ein Client eine Verbindung öffnen will, muß er nicht nur die IP-Adresse des Zielrechners angeben, sondern auch die Portnummer des Servers.

Er bekommt auf dem Clientrechner eine Portnummer zugewiesen.

- Es kann gleichzeitig mehrere TCP-Verbindungen von/zum gleichen Rechner geben.

Verbindungen sind eindeutig identifiziert über vier Komponenten: Die beiden IP-Adressen und die beiden Portnummern. Daher ist auch eine bestimmte Portnummer auf dem Server-Rechner nicht exklusiv für die Dauer der Verbindung belegt (die Verbindungen können ja noch über die Client-Daten auseinander gehalten werden). Viele Server bestehen aus mehreren parallelen Prozessen, von denen sich jeder um eine gerade offene Verbindung kümmert.

# TCP (4)

- Bekannte Portnummern verschiedener Dienste sind:

Port	Service	Port	Service
7	echo	23	telnet
13	daytime	25	SMTP/mail
20	ftp-data	37	time
21	ftp	79	finger
22	ssh	80	WWW

Siehe `/etc/services`, `/etc/inet/services` (unter UNIX). Unter UNIX sind Portnummern unter 1024 für privilegierte Prozesse reserviert.

- Man kann mit dem Programm `telnet` eine TCP-Verbindung öffnen, z.B. `telnet localhost 13`.

# TCP (5)

- TCP ist zuverlässig: Der Empfänger sendet eine Quittung für die empfangenen Daten.

Wenn der Sender die Quittung nicht innerhalb einer bestimmten Zeit erhält, schickt er die Daten erneut. Wenn das nach mehreren Versuchen nicht klappt, wird das Benutzerprogramm informiert.

- Der Aufbau und Abbau der Verbindung erzeugt einen gewissen Overhead.
- Für Dienste, die nur Datenpakete (unzuverlässig) verschicken müssen, ist das Protokoll UDP gedacht (ähnlich IP mit Portnummern).

# HTTP

- HTTP, das “HyperText Transfer Protocol”, wird in einem eigenen Kapitel besprochen.
- Wesentliche Aufgaben sind:
  - ◇ Übermittlung von Metadaten zu einem Datenpaket (z.B. Medientyp, Sprache, Zeichensatz).
  - ◇ Übermittlung von Wünschen des Benutzers (z.B. Sprache).
  - ◇ Ggf. Weiterleitung an andere Adresse.
  - ◇ Unterstützung von Proxies (Zwischenspeichern).
  - ◇ Ggf. Passwortschutz.

# ARP: Address Resolution

- Wenn Rechner  $A$  per IP Daten an Rechner  $B$  im lokalen Netz schicken will, kennt er die IP-Adresse von  $B$ , aber zunächst nicht die Ethernet-Adresse.

Daß  $B$  im gleichen lokalen Netz ist, kann er durch Vergleich der beiden IP-Adressen unter Berücksichtigung der Subnetz-Maske herausfinden.

- ARP dient zur Bestimmung dieser Adresse.

Im wesentlichen schickt er einen Broadcast (an alle Rechner im lokalen Netz): "Ich (IP-Adresse  $X$ , Ethernet-Adresse  $Y$ ) wüßte gerne die Ethernet-Adresse von dem rechner mit IP-Adresse  $Z$ ." Der Rechner mit IP-Adresse  $Z$  antwortet ihm dann mit seinen Daten. Damit nicht jedesmal neu gefragt werden muß, werden die Daten für eine gewisse Zeit lokal zwischengespeichert.

# Point-to-Point Protocol

- PPP ist ein häufig verwendetes Protokoll für die Verbindung zweier Rechner über ein serielles Kabel (ggf. mit Modems und einer Telefonverbindung dazwischen).

Wenn man sich mit einem Modem bei einem ISP (Internet Service Provider) einwählt, verwendet man sehr wahrscheinlich PPP.

- Es ersetzt die Ethernet-Ebene.

Außer dem reinen Austausch von Datenpaketen definiert es auch ein "Link Control Protocol" um die Verbindung auf "Data Link" Ebene aufzubauen, zu konfigurieren, und zu testen. Außerdem gibt es verschiedene "Network Control Protocols" (NCP), um Parameter für den Austausch von Datenpaketen auf höherer Ebene (z.B. IP) abzusprechen (etwa Komprimierung).

# Modems (1)

- “Modulator–Demodulator” : Verwendet die normale analoge Telefon-Verbindung
- V.90 standard: Max. Empfangsrate: 56 Kbit/s, max. Senderate: 33.6 Kbit/s.

V.92 standard (Sommer 2000): 48 Kbit/s Senderate.

- Modems verwenden Datenkompression, um höhere Transferraten zu erreichen.

## Modems (2)

- In Deutschland gibt es unterschiedliche Kabel für Telefone und für Modems.
  - “F” kodierter TAE-Stecker für Telefon ( “Fernsprechen” ), “N” kodiert für Zusatzgeräte wie Modem, Fax, etc. ( “Nicht-Fernsprechen” ).
- Verbindung mit dem Computer über Serielle Schnittstelle (9 pin or 25 pin Verbindung) oder USB.

## ISDN, DSL

- ISDN bringt 64 Kbit/s, und mit 2 Kanälen werden es 128 Kbit/s.

Wenn man zwei Kanäle braucht, muss man auch für zwei zahlen.

- DSL: Über 24 MBit/s (ADSL2+).
- Beispiel: Download Zeit für Rational Rose (234 MB):

Modem	(56 Kbit/s):	10 h, 18 min.
ISDN	(128 Kbit/s):	4 h, 33 min.
DSL	(768 kbit/s):	46 min.

# Inhalt

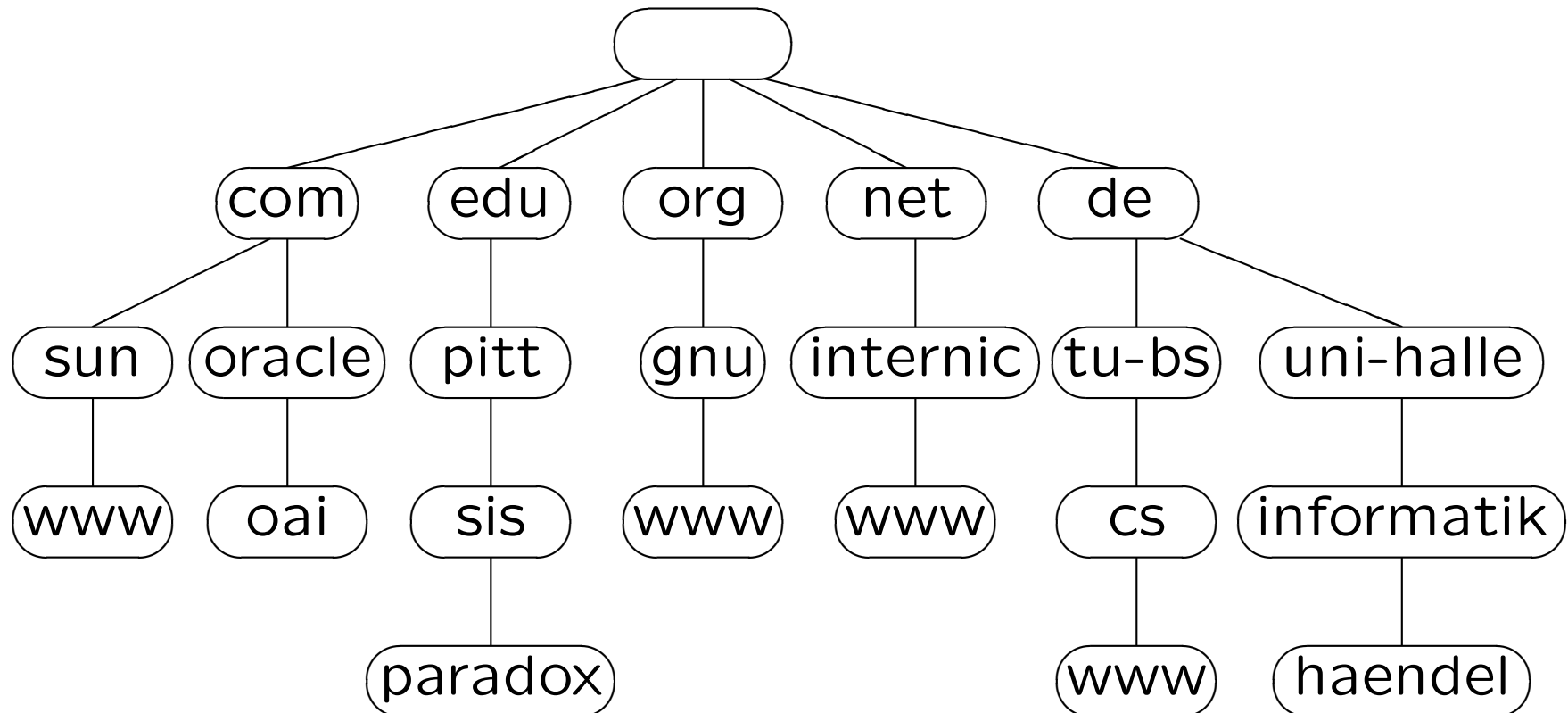
1. Das Internet
2. Geschichte von Internet und WWW
3. Protokoll Schichten ( "Protocol Stack" )
4. Domain Name System
5. Internet Hosting (Eigene Domain)

# Grundlagen (1)

- IP-Adressen sind 32-bit Zahlen. Für uns ist es jedoch einfacher, mit Namen zu arbeiten.
- Das Domain Name System (DNS) ist eine verteilte Datenbank, mit der die Namen und Zahlen verwaltet werden (und andere Informationen).
- Die meisten Internet-Anwendungen brauchen das DNS zur Übersetzung der Hostnames in IP-Adressen.

## Grundlagen (2)

- Die Namen des DNS sind hierarchisch strukturiert:



## Grundlagen (3)

- Am Anfang waren Host-Namen Zeichenketten ohne weitere Struktur ( “flat namespace” ).
- Eine Datei mit der “Name → Zahl” Abbildung wurde verwaltet von InterNIC. Jeder Computer erhielt eine Kopie.
- Diese Lösung funktioniert aber nur mit relativ wenigen Rechnern („nicht skalierbar“):
  - ◇ Die Administration wurde zu kompliziert, weil immer häufiger Änderungen nötig waren.
  - ◇ Immer mehr Namenskonflikte traten auf.
  - ◇ Aufwand (Netzwerk-Last) für Verteilung der Dateien war sehr hoch.

## Grundlagen (4)

- Lösung: Hierarchische Namen
- Eine einfache Baumstruktur wie z. B. beim Unix-Dateisystem
- Im Gegensatz zu Dateinamen schreibt man Hostnamen von den Blättern zur Wurzel (geteilt durch Punkte),  
z.B. `“haendel.informatik.uni-halle.de.”`.
- Ein vollständiger Domain-Name endet mit einem Punkt.

## Grundlagen (5)

- Verteilte Datenbank: kein Computer hat die gesamte Übersicht. Verteilte Administration (Delegation der Teilbäume).
- Jeder Name eines Baumteils wird “Label” genannt.
- Solche Labels können bis zu 63 Zeichen lang werden. Der gesamte Domain-Name darf nicht länger als 255 Zeichen sein.
- Der Vergleich mit dem Domain-Namen ist nicht case-sensitiv.

## Grundlagen (6)

- “Top-Level Domains” (TLDs) sind Kinder des Wurzelknotens, z.B.

- ◇ com, edu, gov, int, mil, nato, net, org  
(“generic domains”)

Die Domains com, org und net können auch außerhalb der USA genutzt werden - sie waren am Anfang für US-Organisation vorgesehen.

- ◇ de, at, ch, uk, fr, ca, jp (“Geographische Domains”).

Die geographischen Domains haben den Ländercode aus zwei Zeichen, definiert in ISO 3166.

## Grundlagen (7)

- Neue allgemeine Domains wurden von ICANN im November 2000 festgelegt:  
aero, biz, coop, info, museum, name.
- Einige Ländercodes werden auch von Firmen, die nicht im Land ansässig sind, verwendet, z. B. .to (Tonga), .tm (Turkmenistan), .tv (Tuvalu), .ag (Antigua).
- Die Blattknoten sind normalerweise Computernamen, die inneren Knoten sind Domainnamen.

Die DNS setzt das nicht durch. Es gibt viele Ausnahmen.

## Grundlagen (8)

- Jeder Knoten kann unterschiedliche Typen von Informationen enthalten. (“resource records”).
- Über 20 unterschiedliche Satztypen sind definiert. Die wichtigsten sind:
  - ◇ A (Address): IP-Nummer.
  - ◇ PTR (Pointer): Name des Computers.
  - ◇ CNAME (Canonical Name): Reale Host-Namen.

verwendet für Aliases. z. B. unter `www.informatik.uni-giessen.de` referenziert der CNAME den Namen `odin.informatik.uni-giessen.de`.

## Grundlagen(9)

- Weitere Typen von DNS Angaben:
  - ◇ HINFO (Host Information): Zwei Zeichenketten zur Beschreibung von Hardware und Betriebssystem.  
Sie sind optional und oft nicht definiert.
  - ◇ MX (Mail Exchanger): Mailserver für diese Domain
  - ◇ NS (Name Server): Dieser Computer verwaltet die DNS-Einträge der Domain.

## DNS als verteilte DB (1)

- Der DNS-Baum ist geteilt in “Zonen”: Bereiche, die von einer bestimmten Stelle verwaltet werden.

Eine Zone ist ein Unterbaum (subtree) des DNS-Baums, aber ohne die Teilbäume, die selbst wieder eine eigene Zone sind.

- z.B. unser Rechenzentrum verwaltet die Zonen `uni-halle.de` und `48.141.in-addr.arpa`.

## DNS als verteilte DB (2)

- Für jede Zone, gibt es mindestens zwei Nameserver (zum Schutz gegen Ausfälle).

z.B. die Wurzelzone hat 13 Nameserver, .de hat 7.

- Ein Server kann für mehrere Zonen verantwortlich sein.
- Der Knoten `uni-halle.de` ist tatsächlich zweimal vertreten: Einmal über die Nameserver der Zone "de" (mit NS Record), und einmal über den Nameserver der eigenen Zone (SOA Record)

Auf diese Art sind die Nameserver miteinander "verbunden".

# Anfrage Ausführung (1)

- Das Modul, das die DNS Anfragen ausführt, wird “Resolver” genannt.

Unter UNIX ist es eine Bibliothek, die von allen Programmen, die das DNS benötigen, verwendet wird.

- Die Resolver müssen mindestens einen DNS-Server kennen, dem sie die Anfrage senden können.

Unter UNIX enthält die Datei `/etc/resolv.conf` Adressen von DNS-Servern. Es gibt auch andere Quellen für Zuordnung von DNS-Namen zu IP-Adressen, z.B. die Datei `/etc/hosts` unter UNIX bzw. `C:\WINDOWS\system32\drivers\etc\hosts` unter Windows.

## Anfrage Ausführung (2)

- Beispiel für `/etc/resolv.conf`:

```
domain informatik.uni-halle.de
nameserver 141.48.3.3
nameserver 141.48.3.51
nameserver 192.76.176.9
```

- Bis zu drei Nameserver können konfiguriert werden.

Wenn keine Antwort von dem ersten kommen sollte, wird der zweite versucht und so weiter. Natürlich müssen IP-Nummern angegeben werden, nicht Namen, da ohne Nameserver ja gerade keine Abbildung von Namen auf Nummern möglich ist. Die Nameserver im Beispiel sind: `ns1.uni-halle.de`, `ns3.uni-halle.de`, `deneb.dfn.de`.

## Anfrage Ausführung (3)

- Der `domain` Eintrag wird verwendet, um lokale Namen wie "haendel" in absolute Namen umzusetzen.
- Für eine Modem-Verbindung via PPP werden die Nameserver in der Regel automatisch zugewiesen.

Unter Windows, kann der Nameserver explizit in einer Dialogbox spezifiziert werden: unter "Systemsteuerung→Netzwerkverbindungen→LAN-Verbindung→Eigenschaften→Internetprotokoll (TCP/IP)→Eigenschaften"

## Anfrage Ausführung (4)

- Beispiel: Wir brauchen eine IP-Adresse von `paradox.sis.pitt.edu`.
- Der Resolver sendet die Anfrage an den ersten konfigurierten Nameserver (`141.48.3.3`, `ns1.uni-halle.de`) via UDP. Der Nameserver erwartet Anfragen am Port 53.

UDP (user datagram protocol) hat weniger Overhead als TCP.

# Anfrage Ausführung (5)

- Es gibt zwei Arten von Anfragen:
  - ◇ Rekursiv: Der Nameserver wird versuchen, die Lösung komplett zu bringen und kontaktiert einen anderen Nameserver, falls es nötig ist.

Das ist der Normalfall für DNS-Clients (Resolver).

- ◇ Iterativ: Wenn der Nameserver die Antwort auf die Anfrage nicht findet, sendet er einen Verweis auf einen Nameserver, der als nächstes gefragt werden sollte.

Das ist der Normalfall für Anfragen zwischen Nameservern.

## Anfrage Ausführung (6)

- Der Nameserver unseres Rechenzentrums weiß keine Antwort auf die Anfrage, aber er kennt die Adresse des Root-Servers des DNS.

siehe auch: [ftp://ftp.rs.internic.net/domain/named.root].

Name	IP-Number	Original Name / Comments
A.ROOT-SERVERS.NET.	198.41.0.4	NS.INTERNIC.NET
B.ROOT-SERVERS.NET.	128.9.0.107	NS1.ISI.EDU
...	...	...

- Der Nameserver unseres Rechenzentrums sendet die Anfrage an einen Root-Server, z.B. A.ROOT-SERVERS.NET

## Anfrage Ausführung (7)

- Dieser Rechner ist zufällig nicht nur für die Root-Zone verantwortlich, sondern auch für die Top-Level-Domain `edu`.
- Er kennt also die Nameserver der Domain `pitt.edu` (die delegiert wurde, d.h. eine andere Zone):  
`ns0-qip.ns.pitt.edu`, `ns1-qip.ns.pitt.edu`
- Der Rootserver sendet jetzt Namen und IP-Adressen von diesen beiden Nameservern zurück zum Nameserver unseres Rechenzentrums.

## Anfrage Ausführung (8)

- Der Nameserver unseres Rechenzentrums sendet die Anfrage jetzt zu `ns0-qip.ns.pitt.edu`.

Wir gehen davon aus, dass hier eine iterative Abfrage verwendet wurde. Der Root-Nameserver antwortet normalerweise nicht auf rekursive Abfragen, aber auch andere Nameserver akzeptieren sie häufig nicht.

- Er ist zuständig für `pitt.edu`, aber die Domain `sis.pitt.edu` ist wieder in einer anderen Zone.
- Also sendet `ns0-qip.ns.pitt.edu` jetzt die für `sis.pitt.edu` verantwortlichen Nameserver zurück zu unserem Nameserver: `icarus.lis.pitt.edu`, `acheron.lis.pitt.edu`, `thing.cs.pitt.edu`.

## Anfrage Ausführung (9)

- Als nächstes sendet der Nameserver unseres Rechenzentrums die Anfrage zu `icarus.lis.pitt.edu`.
- Von ihm kommt die abschließende Antwort: die IP-Adresse von `paradox.sis.pitt.edu` ist `136.142.116.28`.
- Insgesamt arbeiteten vier DNS Server an der Anfrage.
- Im Hinblick auf die Reduzierung des Netzwerkverkehrs haben alle DNS-Server einen Puffer (Cache) für "Resource-Records", die sie vor kurzem ermittelt haben.

# Anfrage Ausführung (10)

- Der Administrator einer Zone kann definieren, wie lange der “Resource Record” dieser Zone gepuffert werden soll.
- z. B. wenn der Nameserver unseres Rechenzentrums wieder nach der Adresse `paradox.sis.pitt.edu` gefragt wird, kommt die Antwort direkt aus dem Cache.

Dem Client wird mitgeteilt, dass es sich hierbei um eine “non-authoritative answer” handelt und welcher Nameserver die “authoritative answer” hat.

# nslookup (1)

- Unter UNIX kann das Programm nslookup für Anfragen an das DNS verwendet werden.

Es existiert auch unter Windows 2000, aber nicht unter Windows 98 oder Windows ME. Es arbeitet im interaktiven Modus (ohne Parameter aufgerufen) oder im nicht-interaktiven Modus (wenn die Anfrage bereits auf der Kommando-Zeile erfolgte). Das Programm kann mit "exit" verlassen werden. Einen Überblick über die Kommandos gibt "help".

- nslookup haendel.informatik.uni-halle.de

Server: ns1.Uni-Halle.DE

Address: 141.48.3.3

Name: haendel.informatik.uni-halle.de

Address: 141.48.14.50

## nslookup (2)

- nslookup 141.48.14.50  
Server: ns1.Uni-Halle.DE  
Address: 141.48.3.3  
Name: haendel.informatik.Uni-Halle.DE  
Address: 141.48.14.50
- nslookup paradox.sis.pitt.edu  
Server: ns1.Uni-Halle.DE  
Address: 141.48.3.3  
Non-authoritative answer:  
Name: paradox.sis.pitt.edu  
Address: 136.142.116.28

## nslookup (3)

- Man kann auch einen spezifischen Nameserver wählen

```
nslookup paradox.... acheron.sis.pitt.edu
```

```
Server: acheron.sis.pitt.edu
```

```
Address: 136.142.116.73
```

```
Name: paradox.sis.pitt.edu
```

```
Address: 136.142.116.28
```

- Im interaktiven Modus wird der Server gewählt mit:

```
NSLOOKUP> server acheron.sis.pitt.edu.
```

- Dann kann man Anfragen eingeben, z.B.

```
NSLOOKUP> paradox.sis.pitt.edu.
```

## nslookup (4)

- Man kann angeben, welche Einträge ausgegeben werden sollen.

z.B. mit `-q=mx` bekommt man den Mailserver

```
>nslookup -q=mx uni-halle.de
Server: ns1.uni-halle.de
Address: 141.48.3.3

uni-halle.de MX preference = 100, mail exchanger = mailgate2.urz.uni-halle.de
uni-halle.de MX preference = 10, mail exchanger = mailgate4.urz.uni-halle.de
uni-halle.de MX preference = 50, mail exchanger = mailgate.urz.uni-halle.de
uni-halle.de nameserver = ns1.uni-halle.de
uni-halle.de nameserver = deneb.dfn.de
uni-halle.de nameserver = ns3.uni-halle.de
uni-halle.de nameserver = ns2.uni-halle.de
mailgate.urz.uni-halle.de internet address = 141.48.3.51
mailgate2.urz.uni-halle.de internet address = 141.48.3.8
mailgate4.urz.uni-halle.de internet address = 141.48.3.25
ns1.uni-halle.de internet address = 141.48.3.3
...
```

## nslookup (5)

- Einige Nameserver erlauben es, alle Einträge in ihrer Domain aufzulisten:

```
UNIX> nslookup
```

```
> server regulus.informatik.uni-hannover.de.
```

```
> ls informatik.uni-hannover.de.
```

```
> exit
```

- Leider haben viele Administratoren diese Möglichkeit ausgeschaltet.
- Mit der Option `d2` werden Anfrage und Antwort komplett detailliert aufgelistet.

# Inhalt

1. Das Internet
2. Geschichte von Internet und WWW
3. Protokoll Schichten ( "Protocol Stack" )
4. Domain Name System
5. Internet Hosting (Eigene Domain)

# Domain Registration (1)

- Die Nameserver der Zone “de” werden von der DENIC eG [<http://www.denic.de>] verwaltet.

Die Mitglieder der DENIC Genossenschaft sind die Internet Service Provider in Deutschland.

- Für den eigenen Nameserver (oder für einen für die eigene Domain unter de verantwortlichen Nameserver) kann die Delegation der Domain zu diesem Server bei der DENIC beantragt werden.

- DENIC speichert NS-Records für diese Domains auf ihren Servern
- Kosten: 116 Euro für das erste Jahre, 58 Euro für die folgenden
- Gesonderte Bedingungen für DENIC-Mitglieder

## Domain Registration (2)

- Wenn man keinen eigenen Nameserver hat, kann man beantragen, dass die (A und MX)-Records direkt auf einem DENIC-Nameserver gespeichert werden.

Der Preis (für 5 Einträge) entspricht dem Preis für die Delegation der Domain. Registriert man die Domain (und Webspace) bei einem DENIC-Mitglied, erhält man Domain und Einträge wesentlich günstiger

- Die DENIC legt auch die Regeln für gültige Domainnamen unter `de` fest:

z.B. mind. drei Zeichen (mind. ein Buchstabe), Maximallänge 63 Zeichen, keine Namen von TLDs, keine Abkürzungen von Kfz-Kennzeichen, kein Bindestrich an der ersten, letzten, dritten und vierten Position.

## Domain Registration (3)

- Die ICANN (Internet Corporation for Assigned Names and Numbers) ist verantwortlich für Top Level Domains [<http://www.icann.org/>].

Für jede TLD werden von der ICANN mehrere Registrierungsorganisation (Registrars) akkreditiert.

- Domains unterhalb der allgemeinen TLDs wurden von 1993 bis 1998 nur von der Network Solutions, Inc. [<http://www.networksolutions.com>] vergeben.

Formal war diese das InterNIC.

# Domain Registration (4)

- Heute gibt es viele Registrars für Domains unterhalb der allgemeinen TLDs com, org, net :
  - ◇ [<http://www.icann.org/registrars/accredited-list.html>]
  - ◇ [<http://www.internic.net>]
- Meistens gibt es nur einen Registrar für eine Domain.

DENIC ist die einzige für die Domain de, aber viele Firmen und Organisation sind Mitglieder in der DENIC.
- IP-Nummern und `inaddr.arpa`-Domains werden in Europa von RIPE NCC verwaltet.

# Domain Registration (5)

- Domain Registrars wie DENIC müssen die folgenden Information erfassen:
  - ◇ mindestens zwei Nameserver
  - ◇ den Besitzer der Domain (z.B. Firma),
  - ◇ Kontakt für Verwaltung,
    - z.B. eine Person, die Entscheidungen zur Domain treffen kann.
  - ◇ Rechnungsadresse,
  - ◇ technischer Ansprechpartner,
  - ◇ Administrator der Zone (Nameserver).

# Domain Registration (6)

- Die erfassten Daten können mit dem `whois`-Kommando bzw. einem Whois Web Interface abgefragt werden.
- Jeder Registrar hat seine eigene Datenbank - man muss die jeweils passende auswählen.
  - ◇ [<http://www.denic.de/servlet/Whois>]  
Die DENIC `whois` DB enthält Informationen über `.de` Domains.  
z.B. `whois -h whois.denic.de <Domain>`.
  - ◇ [<http://www.internic.net/whois.html>]  
Die Internic DB contains information enthält Informationen über `.com`, `.net`, `.org`, `.edu`, aber nur Verweise auf die DB des Registrars.
  - ◇ “Universal Whois” : [<http://www.uwhois.com>].

# Internet Hosting (1)

- Wenn man mit dem Internet verbunden ist, kann man seinen eigenen Webserver betreiben.

Natürlich sollte man auch rund um die Uhr verbunden sein, um ohne Einschränkungen erreichbar zu sein. Eine "normale" Flatrate ist dafür nicht gedacht z.B. trennt die Telekom die Verbindung automatisch nach 24 Stunden. Man kann sich sofort wieder verbinden aber man erhält dann eine neue IP Adresse. Es gibt Dynamische DNS Server, mit denen sich die Abbildung Name → IP-Adresse einfach ändern lässt. So kann man den Domain-Namen unabhängig von der IP behalten.

- Man kann auch Speicherplatz auf einem Webserver mieten (WWW Hosting).

Viele Online-Angebote und ISP Tarife enthalten Speicherplatz, aber nur in einem Unterverzeichnis des Webserver des Anbieters.

## Internet Hosting (2)

- Man kann auch eine Domain registrieren und von dieser Domain auf einen Server, auf dem man Speicherplatz gemietet hat, weiterleiten.
  - ◇ Meistens werden viele Domains auf denselben Host weitergeleitet.
  - ◇ Für private Zwecke wäre es zu teuer, wenn man einen kompletten Server mietet.
- Wenn der Server mehrere IP-Adressen hat, kann er die Webseiten abhängig von der jeweils verwendeten Adresse ausliefern.

Obwohl es nur einen Computer mit nur einem Webserver gibt, wirkt er von außen wie mehrere Computer mit viele Webservern.

## Internet Hosting (3)

- IP-Adressen sind eine knappe Resource. Deshalb werden viele Domains auf dieselbe Adresse abgebildet.
- In HTTP/1.1 ist der Name des Servers in der Anfrage enthalten. Ein einzelner Computer kann so Server für viele Domains sein.

Mit HTTP/1.0 funktioniert das nicht, da der Webserver keine Angaben zum angefragten Hostname erhält. Man hatte angenommen, dass der WWW-Server seinen eigenen Hostname kennt.

# Selection Criteria (1)

- Will one be registered as domain owner and administrative contact for the domain at DENIC?

Important for changing to a different web hosting company later. Even if one is the domain owner, there may be a fee for the domain transfer.

- Can arbitrary web pages and other files be stored?

Some cheap offers have only one webpage with a fixed format. Otherwise, one gets FTP access to upload arbitrary files on the web server.

- How many domains? Also .com, .net, .org, etc?  
Subdomains?

## Selection Criteria (2)

- How much disk space for files on the web server (“webspaces”)?

- Data transfer volume per month?

If the pages are accessed very often (e.g. from a robot operated by some hacker), that can cause large extra costs. Some companies offer to simply switch the website off when the quota is reached, others only send a warning email, others maybe not even that.

- Does one get statistical data about web accesses?

Nice graphical representation? How detailed (e.g. server log files)?

- Redirection to an existing web server?

## Selection Criteria (3)

- Can one use CGI programs?

Only select one from a fixed collection or write one's own programs? If yes, what languages are supported? Telnet/SSH access would be useful for debugging the programs, but is not strictly needed. The operating system might be interesting.

- Can the CGI programs access a database?

If yes, what DBMS? Are there restrictions in the SQL language (e.g. MySQL)? Does it have support for transactions? What about the safety of the data: Does the DBMS write log files, does anybody make backup copies?

- Online shop included?

## Selection Criteria (4)

- SSL-support (https:)?

Passwords and credit card numbers should be encrypted while sent through the internet.

- Can certian pages be password-protected?

- Email Accounts?

POP3 mailboxes, email forwarding, autoresponder? How large can the emails and their attachments be? How large is the mailbox? Are emails automatically deleted after some time? Can an SMS be sent for incoming emails (and what does this cost)? Can emails be sent via fax?

## Selection Criteria (5)

- WAP-support?
- Is software included?

Some software is shipped on a CD, other software has to be used via the web on the server of the web hosting company. Sometimes the software is much more expensive if bought separately.

- Technical support in case of problems?

Sometimes this is very expensive (e.g. telephone support via 0190-number). Is there a hotline 7 days a week, 24 h a day?

## Selection Criteria (6)

- How available is the web server?

If the web server is often not reachable/down, this is a problem. E.g. are there redundant internet connections? Do they have battery backup for power failures?

- How fast is the internet connection of the server?

If most customers live in Germany, the server should probably be in Germany with a fast connection to the DE-CIX.

- Does the web hosting company make backups of the data on the web server? Do they use RAID-systems?

# Domain Laws (1)

- See: [<http://www.domain-recht.de/>]
- When one registers a domain, one has to sign that one does not violate rights of other persons or companies, and that one will pay the fees for lawsuits.

This is probably the main reason why one is today normally registered as domain owner.

- Trademarks/company names may not be registered, even if by chance it is ones own family name.

Also combinations like `microsoft-haters.de` can be problematic as well as small changes like `microsaft.de`.

## Domain Laws (2)

- Names of persons belong to these persons.
- Titles of journals, books, software, films are protected if they are very well known or there is a possibility of confusion.
- Names of cities or countries belong to them.  
In addition, DENIC does not permit auto license city codes.
- Names of government agencies belong to them.
- “Typing error domains” like `aliavista.com` belong to the owner of the corresponding well known domain `altavista.com`.

## Domain Laws (3)

- General descriptive names like `database-course.de` are not forbidden.

Such domains gets the first person who requests them, even though there are other database courses.

- However, `mitwohnzentrale.de` was successfully sued.

The appeal is still open and some experts feel that the judge made an error. This would be a problem for many domains, e.g. also `buch.de`.

- Domains are sold, sometimes for large sums.

The internet community thinks that this is an abuse. Judges assumed for quite some time that trading domains is immoral. The domain `loans.com` was sold for 3 Mio \$.

## Domain Laws (4)

- The rules for domain names depend on the top-level domain (they are determined by the registrar).

At least, the registrar has the technical possibility to change the name server entries. Of course, one can also go before a usual court.

- Some domains were lost because bills of the registrar were not paid in time.
- All registrars for the domains .com, .org, .net have adopted the “Uniform Domain Name Dispute Resolution Policy” .

[<http://www.icann.org/udrp/udrp-rules-24oct99.htm>]

## Domain Laws (5)

- The UDRP requires that a domain is transferred if the complainant proves three things:
  - ◇ The domain name is identical or confusingly similar to a trademark owned by the complainant.
  - ◇ The current owner of the domain has no rights or legitimate interests in the domain.
  - ◇ The domain was registered with evil purpose.

E.g. in order to sell it, to get page hits because of the confusion, or to disturb the business of the complainant.