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`
- T. Socolofsky, C. Kale: A TCP/IP Tutorial. RFC 1180.

- Dank an Leonhard Knauff vom URZ für viele Erklärungen!

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

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".

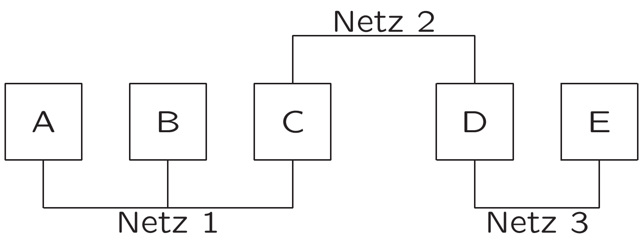
- Im Januar 2008 waren ca. 542 Millionen Rechner an das Internet angeschlossen

  Quelle: [http://www.isc.org/ds/].
  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 13-13.

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.
- Spezielle Adressen:
  - ◇ Alle Bits des Host-Anteils 0: Netzwerk.
  - ◇ Alle Bits des Host-Anteils 1: Broadcast.

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)

- Man kann aus der `whois` Datenbank u.a. herausfinden, wer für eine IP-Adresse zuständig ist:

```
UNIX> whois -h whois.ripe.net 141.48.0.0
http://www.ripe.net/perl/whois
```

  Es sind nur die Netzwerke eingetragen, nicht die einzelnen Rechner (d.h. der Rechner-Anteil besteht nur aus 0-Bits). Alternativ kann man die IP-Nummer auch in einen Domain-Namen umwandeln, und dann nach Informationen zu der Domain suchen (s.u.).

- Es gibt offiziell freie IP-Nummern, z.B. `192.168.x.y`.

  Verwendung z.B. wenn alle Kommunikation nach außen über einen Proxy erfolgt, so daß die lokalen IP-Nummern niemals nach außen gelangen.

Internet Adressen (7)

- 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 (8)

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

  Dan 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 Kernnetzknoten 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 Wochenendes). 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/de/beratung/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 January 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 nordamerikanischen Forschungsnetzen,
  - ◇ via DE-CIX zu kommerziellen Internetprovidern 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)

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

  Man kann dieses Kommado nicht nur vom eigenen Rechner aus verwenden: [http://www.traceroute.org].
- Beispiel: `traceroute 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	7200vxr.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 `Crtl`+`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 Rechernamen 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)
- `arp wega` zeigt die Ethernet-Adresse (MAC Adresse) von `wega` an, `arp -a` alle im Moment gespeicherten.

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

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 kooridiniert 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 73 500.

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).

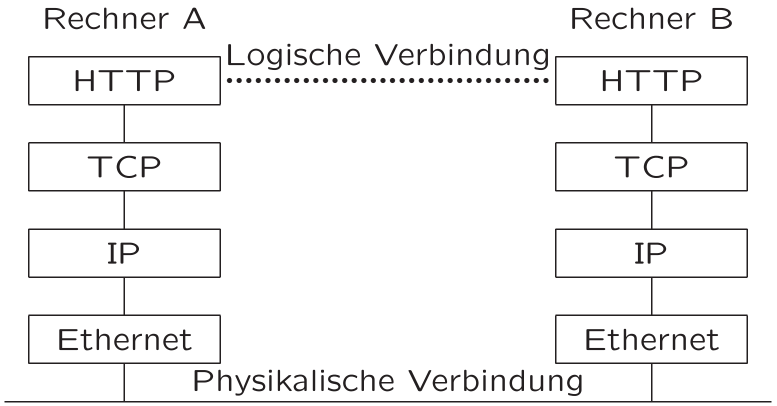
  Dies war die erste Version, die formal über eine SGML DTD definiert war. RFCs ("Request for Comments") sind Internet Standards (und andere Internet Dokumente) siehe z.B. [http://www.rfc-editor.org].

Inhalt

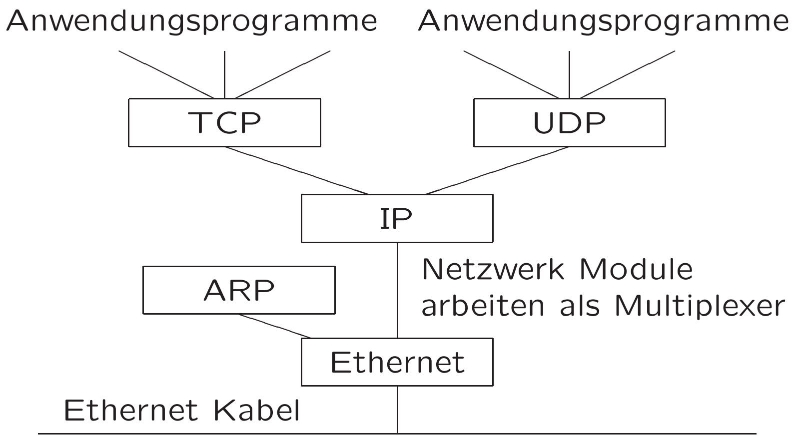
Schichtenarchitektur

- 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.

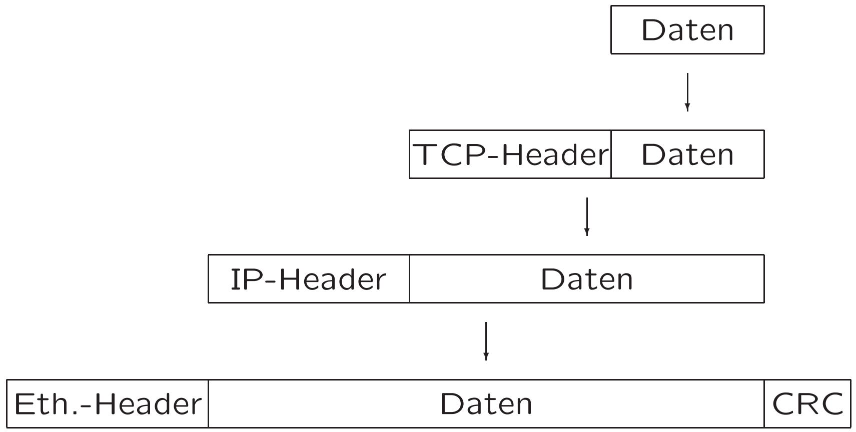
"Protocol Stack"



Protokolle auf einem Rechner



Verpacken von Datenpaketen



Vergleich mit OSI-Model

OSI	TCP/IP	
Application	Application	Telnet, FTP, HTTP, etc.
Presentation		
Session		
Transport	Transport	TCP, UDP
Network	Network	IP, ICMP
Data Link	Link	Device Driver Hardware
Physical		

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 Ω-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.

IPv6 (3)

- IPv6 Pakete können von einem IPv6 Netz in ein anderes IPv6 Netz durch Tunnelung über das IPv4 Internet verschickt werden.

  Tunnelung ist eine allgemeine Technik, um ein Protokoll über ein anderes Protokoll zu befördern. Am Anfang des Tunnels werden IPv6 Pakete in einen IPv4 “Umschlag” eingepackt (d.h. das IPv6 Paket sind die Nutzerdaten des IPv4 Paketes, das dann einen IPv4 Header bekommt). Am Ende des Tunnels werden sie ausgepackt, d.h. der IPv4 Header wird wieder entfernt, und die Nutzerdaten als IPv6 Paket in das andere Netz eingespeist.
- Bisher wird IPv6 nur minimal genutzt, aber man schätzt, daß 2011 der letzte 8-Bit Block von IPv4 Adressen vergeben wird.

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 Verbidnung 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 der 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:
  - ◇ Übermitlung 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": Uses a normal analog telephone connection.
- V.90 standard: Max. receive rate: 56 Kbit/s, max. send rate: 33.6 Kbit/s.

  V.92 standard (Summer 2000): Now 48 Kbit/s send rate.
- Modems usually do a data compression, thus higher data rates can be reached.

  V42.bis can compress under optimal circumstances to 25%.

Modems (2)

- In Germany, there are different telephone cable connectors for usual telephones and for modems.

  "F" coded TAE-connector for telephone ("Fernsprechen"),
  "N" coded for modem, Fax, etc. ("Nicht-Fernsprechen").
  Today, there is usually an "F" coded connector in the middle, and two "N" coded at both sides. An "F"-coded plug is at the bottom slightly wider, an "N"-coded in the second part from the top.
- Connected to the computer via a serial line (9 pin or 25 pin connector) or USB.
- Usually have the capability to send/receive faxes.

ISDN, DSL

- ISDN gives 64 Kbit/s, but two channels can be coupled to get 128 Kbit/s.

  If one uses two channels, one has to pay for both (i.e. two times the normal rate).
- DSL: Up to 24 MBit/s (ADSL2+).
- E.g. download time for 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.

Configuring a Modem (1)

- Connection data of some German Call-by-Call ISPs:
  - MSN Easysurfer Power [0193670]

    User: `msn@easysurfer-power.de`, Password: `msn`. If one has preselected a different phone company, it might by safer to use the area code 01033 (Telekom) in order not to pay for the phone call. Currently costs 1.17 cent/min. Time is measured in seconds. Uses lines of MCI Worldcom. Nameserver: `192.76.144.66`.
  - Arcor Spar: Internet by Call [0192076]

    Area code 01070 (Arcor), User: `arcor-spar`, Password: `internet`. Currently costs 0.67 cent/min (Mo-Fr 9-18) or 0.79 cent/min (nights and weekends) plus 8.1 cent per connection (counts whole minutes). Nameserver: `145.253.2.11`, `145.253.2.75`.

Configuring a Modem (2)

- Students of the University of Halle can use e.g. 0345/6860 and 0345/55-21940 with their username and network password from the computing center.

  See [http://pc-serv.urz.uni-halle.de/einwahl/index.htm]. The modem must support V.42 or MNP4 error correction.

- Price comparisons: [http://www.teltarif.de], [http://www.heise.de/itarif/].
- It has happened that dialer programs got as "trojan horses" on computers and changed the internet connection to very expensive 0190-numbers.

Configuring a Modem (3)

- Under Properties/Networking one chooses "PPP" as "Type of Dial-Up Server" and "TCP/IP" as allowed network protocol.
- Under Properties/Dailing one can configure automatic disconnection for idle connections.

  See also modem configuration under "Connection".

- A usual modem configuration is 8 data bits, no parity bit, and 1 stop bit.

  The interface speed is usually 115200 bit/s. It should be higher than the actual transfer speed, because the modem compresses the data.

Inhalt

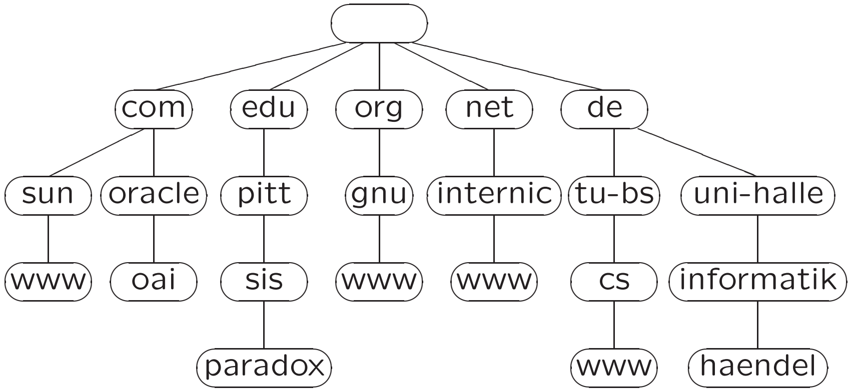
Foundations (1)

- The real IP-addresses are 32-bit numbers. But for humans, it is easier to work with/remember names.
- The Domain Name System (DNS) is a distributed database that manages the mapping from names to numbers (plus other information).
- Most internet applications use the DNS to translate host names into the IP-addresses.

  E.g. Netscape prints "`Looking up host: xyz`" while the DNS-query runs. If it then prints "`Netscape is unable to locate the server xyz`", the name `xyz` is not stored in the DNS (probably a typing error) or no DNS server can be reached (configurarion error, network problem).

Foundations (2)

- Names in the DNS are hierarchically structured:



Foundations (3)

- At the beginning, host names were character strings without further structure ("flat namespace").
- A file with the "Name → Number" map was managed at InterNIC, every computer fetched a copy.
- But this works only with a small number of hosts:
  - ◇ The administration became too complicated because of the more and more frequent changes.
  - ◇ Name conflicts occured more and more often.
  - ◇ The network load for the distribution of the file grows quadratically with the number of hosts.

Foundations (4)

- Solution: Hierarchical namespace.
- Similar tree structure as e.g. the UNIX file system.
- In contrast to file names, host names are written from the leaf towards the root (separated by periods), e.g. "`haendel.informatik.uni-halle.de.`".
- A fully qualified/absolute domain name ends in a period (the name of the root node is empty).

  The completion of other names depends on the DNS-Software. The usual solution is to append the local domain if the name does not contain a period, and otherwise to assume that the name is complete.

Foundations (5)

- If one is currently on "`paradox.sis.pitt.edu`", it normally does not work to specify "`www.math`" to reach "`www.math.pitt.edu`".

  Older software tried to attach portions of the current domain. Suppose one is currently on `a.b.com` and wants to reach `www.pitt.edu`. The software tried "`www.pitt.edu.a.b.com.`", "`www.pitt.edu.b.com.`", "`www.pitt.edu.com.`", and "`www.pitt.edu.`" in this sequence. However, a domain `edu.com` was registered, and it would have been possible to redirect all traffic from `com`-hosts to `edu`-hosts to this domain.
- Distributed database: No computer (DNS server) contains the entire map.
- Distributed administration (delegation of subtrees).

Foundations (6)

- Each name of a node in the tree is called a “label”.
- Such labels can be at most 63 characters long.

  They should start with a letter and otherwise contain only digits and the hyphen “-”. The hyphen should not be the last character (see RFC 952 and RFC 1123). For instance, an underscore “_” is explicitly excluded. However, the DNS software should be able to work with arbitrary (non-empty) binary strings as labels (see RFC 2181). It is being discussed to permit national characters in domain names.
- The whole domain name may not be longer than 255 characters.
- The domain name comparison is not case-sensitive.

Foundations (7)

- “Top-Level Domains” (TLDs) are children of the root node, e.g.
  - ◇ `com`, `edu`, `gov`, `int`, `mil`, `nato`, `net`, `org` (“generic domains”)

    The domains `com`, `org` and `net` can also be used by non-US organizations (at first, they were intended only for US organizations).
  - ◇ `de`, `at`, `ch`, `uk`, `fr`, `ca`, `jp` (“geographic domains”).

    The geographic domains are the two character country codes defined in ISO 3166.
  - ◇ `arpa` (for inverse map, see below).

Foundations (8)

- New generic domains were selected by the ICANN board (see below) in November 2000: `aero`, `biz`, `coop`, `info`, `museum`, `name`.
- Some country codes are also used by companies not really located in the country, e.g. `.to` (Tongo), `.tm` (Turkmenistan), `.tv` (Tuvalu), `.ag` (Antigua).
- Normally the leaf nodes are names of computers whereas inner nodes are proper domain names.

  The DNS does not enforce this and there are many exceptions. In the DNS, every name that is generated by composing the labels from a node towards the root is called a domain name.

Foundations (9)

- Every node can contain different types of information (“resource records”).
- About 20 different record types were defined. The most important are:
  - ◇ `A` (Address): IP-Number.
  - ◇ `PTR` (Pointer): Name of the computer.

    These records are used to map IP-numbers to host names (see below).
  - ◇ `CNAME` (Canonical Name): Real host name.

    Used for aliases. E.g. under `www.informatik.uni-giessen.de` a `CNAME` record exists that refers to `odin.informatik.uni-giessen.de`.

Foundations (10)

- Types of DNS resource records, continued:
  - ◇ `HINFO` (Host Information): Two strings that describe hardware and operating system.

    They are optional, and often not defined.
  - ◇ `MX` (Mail Exchanger): Host that acts as mail server for this domain (or this computer).

    One can specify several computers with different priorities. Smaller numeric values mean that the computer is tried first.
  - ◇ `NS` (Name Server): Computer that manages the DNS records for this domain.

Foundations (11)

- Types of DNS resource records, continued:
  - ◇ `SOA` (Start of Authority): Administrative information for delegated domain (see below).

    Contains the name of the primary name server, the email address of the name server administrator (with "." instead of "@"), a serial number of the domain data, the time interval after which secondary name servers ask the primary name server for updates (REFRESH), the retry time interval (if the primary name server connot be reached), how long secondary name servers may declare the data as authorative if the refresh fails (EXPIRE), and the minimum time-to-live (how long data may be cached).
  - ◇ `AAAA`: IPv6 address.

Foundations (12)

- The mapping of IP numbers to host names is done via a special domain `in-addr.arpa`.
- E.g. in the node

  `50.14.48.141.in-addr.arpa`

  a `PTR`-record with the following contents is stored:

  `haendel.informatik.Uni-Halle.DE`

  The IP-address of this machine is `141.48.14.50`. The DNS always lists labels from the local to the global, i.e. in the opposite sequence.
- Not all computers that are connected to the internet have DNS entries.

DNS as Distributed DB (1)

- The DNS-tree is divided into "zones": Parts that are administered at one place.

  A zone is a subtree of the DNS tree, but without those of its own subtrees that are zones by themselves. Root nodes of zones in the DNS tree are marked by data records of the type `SOA` ("Start of Authority"). The zone mechanism can be compared with disks mounted into a UNIX file system.
- E.g. our computing center administers the zones `uni-halle.de` and `48.141.in-addr.arpa`.

DNS as Distributed DB (2)

- For each zone, there are at least two name servers (in order to protect against failures).

  E.g. the root zone has 13 name servers, `.de` has 7.

- A server can be responsible for multiple zones.
- The node `uni-halle.de` is actually represented twice: Once on the name servers of the zone "de" (with `NS` Record), and once on the name servers of its own zone (with `SOA` Record).

  In this way, the name servers are linked together.

Query Execution (1)

- The module that executes DNS queries is called a "Resolver".

  Under UNIX, it is a library that is linked to programs that use the DNS.

- The resolver must know at least one DNS server, to which it can send the query.

  Under UNIX `/etc/resolv.conf` contains addresses of DNS servers. However, there are also other sources for the mapping from names to IP-numbers, e.g. `/etc/hosts` and the NIS/NIS+ database. Under Solaris, an entry in `/etc/nsswitch.conf` determines, which sources are queried in which sequence.

Query Execution (2)

- Example for `/etc/resolv.conf`:

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

- Up to three name servers can be configured.

  If there should be no answer from the first, the second is tried, and so on. Of course, IP-numbers must be specified, not names. The name servers in the example are: `ns1.uni-halle.de`, `ns3.uni-halle.de`, `deneb.dfn.de`.

- The `domain` entry is used to translate local names like "`haendel`" in absolute names.

Query Execution (3)

- If one uses a modem connection via PPP, name servers are normally automatically assigned

  Under Windows, name servers can also be explicitly specified in the dialog box "Properties→Networking→TCP/IP Settings".

- Example: Suppose we need to know the IP-address of `paradox.sis.pitt.edu`.
- The resolver sends the query to the first configured name server (`i141.48.3.3`, `ns1.uni-halle.de`) via UDP. The name server listens on port 53.

  UDP (user datagram protocol) has less overhead than TCP.

Query Execution (4)

- There are two kinds of queries:
  - ◇ Recursive: The name server is asked to resolve the query completely, and to contact other name servers itself if necessary.

    This is the usual case for DNS-clients (resolvers).
  - ◇ Iterative: If the name server does not know the answer to the query, it sends back a reference to a name server that should be asked next.

    This is the usual case for queries between name servers.

Query Execution (5)

- The name server of our computing center does not know the answer to the query, but it knows the addresses of the root servers of the DNS.

  See [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
C.ROOT-SERVERS.NET.	192.33.4.12	C.PSI.NET
D.ROOT-SERVERS.NET.	128.8.10.90	TERP.UMD.EDU
E.ROOT-SERVERS.NET.	192.203.230.10	NS.NASA.GOV
F.ROOT-SERVERS.NET.	192.5.5.241	NS.ISC.ORG
G.ROOT-SERVERS.NET.	192.112.36.4	NS.NIC.DDN.MIL
H.ROOT-SERVERS.NET.	128.63.2.53	AOS.ARL.ARMY.MIL
I.ROOT-SERVERS.NET.	192.36.148.17	NIC.NORDU.NET
J.ROOT-SERVERS.NET.	198.41.0.10	temporarily housed at NSI (InterNIC)
K.ROOT-SERVERS.NET.	193.0.14.129	housed in LINX, operated by RIPE
L.ROOT-SERVERS.NET.	198.32.64.12	temporarily housed at ISI (IANA)
M.ROOT-SERVERS.NET.	202.12.27.33	housed in Japan, operated by WIDE

Query Execution (6)

- The name server of our computing center sends the query to a root server, e.g. `A.ROOT-SERVERS.NET`.
- This happens to be responsible not only for the root domain, but also for the top level domain `edu`.
- Thus, it knows the name servers for the domain `pitt.edu` (which is delegated, i.e. a different zone): `ns0-qip.ns.pitt.edu`, `ns1-qip.ns.pitt.edu`, ...
- The root server now sends names and IP-numbers of these three name servers back to the name server of our computing center.

Query Execution (7)

- The name server of our computing center sends the query now to `ns0-qip.ns.pitt.edu`.

  We assume here that it uses an iterative query. The root name servers normally do not respond to recursive queries, but other name servers often accept them.
- This is responsible for `pitt.edu`, but the domain `sis.pitt.edu` is again a different zone.
- Thus, `ns0-qip.ns.pitt.edu` now sends the name servers responsible for `sis.pitt.edu` back to our name server: `icarus.lis.pitt.edu`, `acheron.lis.pitt.edu`, `thing.cs.pitt.edu`.

Query Execution (8)

- Next, the name server of our computing center sends the query to `icarus.lis.pitt.edu`.
- From there it finally gets the answer: "The IP-address of `paradox.sis.pitt.edu` is `136.142.116.28`".
- In total, four DNS servers worked on the query.
- In order to reduce the network traffic caused by the DNS, all DNS servers contain a buffer (cache) for "resource records" that they recently received.

Query Execution (9)

- The administrator of a zone can define how long resource records of this zone may be buffered.
- E.g. if the name server of our computing center is asked again for the address of `paradox.sis.pitt.edu`, it returns the answer directly from its cache.

  The client is told that this is a "non-authorative answer", as well as the address of the name server that has the "authorative answer".
- It has happened that name servers returned besides the requested data also falsified "resource records" that were also buffered and used for further queries.

nslookup (1)

- Unter UNIX, the program `nslookup` can be used to query the DNS.

  It also exists under Windows 2000, but not under Windows 98 or Windows ME. There is an interactive mode (if called without parameters) and a non-interactive mode (if the query is already specified on the command line). One can leave the program with "`exit`". An overview of the commands gives "`help`".
- `nslookup haendel.informatik.uni-halle.de`

```
Server: ns1.Uni-Halle.DE
Address: 141.48.3.3
Name: thor.informatik.uni-giessen.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)

- More information about non-authoritive answers:

  `nslookup -q=any paradox.sis.pitt.edu`

```
Server: ns1.Uni-Halle.DE
Address: 141.48.3.3
Non-authoritative answer:
paradox.sis.pitt.edu internet address = 136.142.116.28
Authoritative answers can be found from:
sis.pitt.edu nameserver = acheron.sis.pitt.edu
sis.pitt.edu nameserver = icarus.sis.pitt.edu
sis.pitt.edu nameserver = thing.cs.pitt.edu
acheron.sis.pitt.edu internet address = 136.142.116.73
icarus.sis.pitt.edu internet address = 136.142.116.2
thing.cs.pitt.edu internet address = 136.142.80.5
```


nslookup (4)

- One can also select a specific name server, e.g.

  `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
```

- In interactive mode, the server is selected with

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

- Then one can enter queries to this server, e.g.

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

nslookup (5)

- One can specify which record types are requested:

  `nslookup -q=ns uni-halle.de.`

```
Server: ns1.Uni-Halle.de
Address: 141.48.3.3
uni-halle.de nameserver = ns2.uni-halle.de
uni-halle.de nameserver = ns3.uni-halle.de
uni-halle.de nameserver = deneb.dfn.de
uni-halle.de nameserver = ns1.uni-halle.de
ns1.uni-halle.de internet address = 141.48.3.3
ns2.uni-halle.de internet address = 141.48.3.17
ns3.uni-halle.de internet address = 141.48.3.51
deneb.dfn.de internet address = 192.76.176.9
```


nslookup (6)

- With `-q=any` one gets in addition the `SOA`-record and the `MX`-records ("mail exchanger").

```
...
uni-halle.de preference = 50, mail exchanger = mailgate.urz.uni-halle
uni-halle.de preference = 100, mail exchanger = mailgate2.urz.uni-hal
uni-halle.de
    origin = ns1.uni-halle.de
    mail addr = knauff.urz.uni-halle.de
    serial = 2004101401
    refresh = 10800 (3H)
    retry = 1800 (30M)
    expire = 604800 (1W)
    minimum ttl = 3600 (1H)
uni-halle.de nameserver = ns3.uni-halle.de
uni-halle.de nameserver = deneb.dfn.de
uni-halle.de nameserver = ns1.uni-halle.de
uni-halle.de nameserver = ns2.uni-halle.de
...
```

nslookup (7)

- Some name servers permit to list all entries in their domain:

```
UNIX> nslookup
> server regulus.informatik.uni-hannover.de.
> ls informatik.uni-hannover.de.
> exit
```

- Unfortunately, most administrators recently switched off this possibility.
- With the option `d2` query and answer are listed in complete detail.

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)

- The name servers of the zone "`de`" are managed by the DENIC [http://www.denic.de].

  Members of the DENIC association are internet service providers in Germany.

- If one is connected to the internet and has one's own name servers (or a contract with somebody who operates a name server) one can request the delegation of a domain below `de` to these servers.

  The DENIC then stores `NS`-records for this domain in their name servers. This costs 116 Euro in the first year and 58 Euro in each following year. To their own members, DENIC offers special conditions.

Domain Registration (2)

- If one has no name server, one can request that DNS records (`A` and `MX`) are stored directly in the DENIC name servers.

  The fee for up to 5 records is the same as that for the delegation of a domain. If one rents webspace from a DENIC member, one can get a domain much cheaper, see below.

- The DENIC also defines rules for acceptable domain names below `de`.

  E.g. at least three characters that contain at least one letter, maximal length 63 characters, no names of top level domains, no city codes of car license numbers, no hyphen at the first, last, and third and fourth position.

Domain Registration (3)

- The ICANN (Internet Corporation for Assigned Names and Numbers) decides on top level domains [`http://www.icann.org/`].

  For each top level domain, the ICANN accredits one (or more) registration companies.

- Domains below the generic top level domains were registered from 1993 to 1998 only by Network Solutions, Inc. [`http://www.networksolutions.com`].

  Formally, this was the InterNIC.

Domain Registration (4)

- Now there are many registrars for domains below the generic top level domains `com`, `org`, `net`, see:
  - ◇ [`http://www.icann.org/registrars/accredited-list.html`]
  - ◇ [`http://www.internic.net`]
- Normally there is only one registrar for a domain.

  The DENIC is the only registrar for the domain `de`, but several companies are members in the DENIC.

- IP numbers and `inaddr.arpa`-domains are managed in Europe by the RIPE NCC, see above.

  Until 1996 the DENIC had also this task.

Domain Registration (5)

- Domain registrars like `DENIC` have to collect the following information:
  - ◇ at least two name servers,
  - ◇ the owner of the domain (e.g. company),
  - ◇ an administrative contact,

    A person who may decide things about the domain.

  - ◇ the person responsible for the bill,
  - ◇ a technical contact,
  - ◇ the administrator for the zone (nameserver).

Domain Registration (6)

- The data a registrar must collect can be queried with the `whois` command or a Whois web interface.
- Every registrar has its own database, and one must select the right database, e.g.
  - ◇ [`http://www.denic.de/servlet/Whois`]

    The DENIC `whois` DB contains information about `.de` domains. E.g. `whois -h whois.denic.de` ⟨Domain⟩.

  - ◇ [`http://www.internic.net/whois.html`]

    The Internic database contains information about `.com`, `.net`, `.org`, `.edu`, but only a reference to the database of the registrar.

  - ◇ "Universal Whois": [`http://www.uwhois.com`].

Internet Hosting (1)

- If one is connected to the internet, one can operate one's own web server.

  Of course, one should to be connected around the clock in order for the web server to be reachable without restrictions. A flatrate is not intended for this application, e.g. the Telekom automatically disconnects their flatrate customers after 24 h connection time. One can immediately reconnect, but is assigned a new IP address. There are dynamic DNS servers which permit to change the mapping from names to IP-addresses easily, so that one can keep a stable name.

- An alternative is to rent space on a web server (WWW Hosting).

  Many online services and ISPs include some webspace, but only in a subdirectory of their web server, not under one's own domain.

Internet Hosting (2)

- It is possible to apply for a domain and let this domain name refer to a host on which one has rented web space.
- Often many domains will point to the same host.

  For private applications, it would be too expensive to rent an entire computer.

- If the host has as many IP numbers, it can use the contacted IP address for the decision which web page should be delivered.

  Although it is really only one machine with a single web server, it looks like many machines, each with a web server running.

Internet Hosting (3)

- But IP numbers are a scarce resource. Therefore, today many different domain names are mapped to the same IP number.
- In HTTP/1.1, the request for a web page contains the name of the server, and therefore a single machine can look like many different web servers, although all have the same IP number.

  With HTTP/1.0, it does not work. There the web server does not get the information about the requested host name. The protocol designers assumed that the WWW-server knows its own host name.

Web-Hosting Offers (1)

- 1&1 WebHosting: [`http://www.puretec.de`]
  - ◇ Web Address 4.0: 0.69 Euro/Month.

    All 1&1 prices are plus a 9.60 Euro one-time fee. This package does not include webspace, but one becomes the domain owner. Only `.de` domains can be chosen in this package. One can redirect web page requests to a different URL. One email account is included. The package also includes Email-support, and also telephone support (0800) for the first four weeks.

Web-Hosting Offers (2)

- 1&1 WebHosting, continued:
  - ◇ Basic 4.0: 3.99 Euro/Month.

    If one accepts advertisements on one's web pages, it costs only 2.99 Euro/month. This package includes: 1 domain (`.de`, `.com`, `.net`, `.org`), 25 MB Webspace, FTP access for uploading web files, 10 email accounts, 100 additional aliases, 4000 MB transfer volume/month (each additional MB costs 1.5 cent), Selection from given simple CGI programs, 1&1 WebChat, WAP support, possibility to redirect page requests to a different URL. The package includes NetObjects Fusion MX, PhotoImpact 7 SE, and other software (one has to pay 6 Euro for shipping and handling).

Web-Hosting Offers (3)

- 1&1 WebHosting, continued:
  - ◇ Star Package 4.0: 6.99 Euro/Month.

    Includes: 2 domains, 5 subdomains, domain redirection, 50 email accounts, 500 email aliases, 100 MB webspace, 7500 MB monthly transfer volume, access to apache logfiles, selection from given CGI programs (e.g. counter), access to a graphic library.

Web-Hosting Offers (4)

- 1&1 WebHosting, continued:
  - ◇ Profi Package 4.0: 12.99 Euro/Month.

    Three domains (including `.info`), 10 subdomains, 250 MB webspace, 15 GB transfer volume/month, Server Side Includes (SSI), PHP3, PHP4 (embedded into HTML web pages), MS Frontpage 2002 Extensions, 1&1 WebDatabase.

  - ◇ Premium Package 4.0: 24.99 Euro/Month.

    5 domains (including `.info`, `.biz`), 20 subdomains, 500 MB webspace, 25 GB monthly transfer volume, SSL support (`https:`), arbitrary CGI programs (Perl, Phython, PHP3, PHP4), MySQL database.

Web-Hosting Offers (5)

- 1&1 WebHosting, continued:
  - ◇ Power Package 4.0: 39.99 Euro/Month.

    10 domains (including `.info`, `.biz`), 50 subdomains, 1 GB webspace, 50 GB monthly transfer volume, SSL, CGI, MySQL as before, CRON jobs, SSH access (login on the server).

  - ◇ Own Server: 69 Euro/Month.

    2000 MHz Celeron, 256 MB RAM, 40 GB disk, 75 GB monthly transfer volume. A Linux system with root access costs only 49 Euro/Month. Both require a one-time fee of 99 Euro. Of course, in this way one gets an own IP address.

Web Hosting Offers (6)

- Strato Medien AG: [http://www.strato.de]
  - ◇ Web Visitenkarte A: 0.28 Euro/Month

    Plus 9.90 Euro one-time fee. 1 web page can be constructed with an online form (no FTP access), or the domain can be redirected to another web server. No POP3 mailbox, but one addresses can be forwarded to other accounts. Data transfer 4 GB, each additional MB costs 1.5 Cent.

Web Hosting Offers (7)

- Strato Medien AG, continued:
  - ◇ Web Visitenkarte S: 0.99 Euro/Month

    Plus 19.90 Euro one-time fee. Includes one .de-domain, 2 MB webspace (daily backups), FTP access for uploading web files, domain cannot be redirected, 4 GB traffic, 12 POP3 mailboxes, 120 email aliases, telephone support costs 1.86 Euro/min, no email support, integrated messaging center (Fax/Email, 100 messages/month included), Namo WebEditor 4, Paint Shop Pro 4.15, PromoWare 2.1, GS ShopBuilder Entry, graphics archive.

Web Hosting Offers (8)

- Strato Medien AG, continued:
  - ◇ PowerWeb A: 4.99 Euro/Month.

    3 domains (.de, .com, .net, .org), 85 MB webspace, 85 POP3 mailboxes, 850 email alaises, 7.5 GB data transfer/month (each additional MB costs 0.015 Euro), server statistics, standard CGIs (about 20), GS ShopBuilder Entry 2, online shop CGI without limit on number of articles, WAP homepage, Adobe GoLive 6.0, Adobe Photoshop Elements 2.0, and other software.

  - ◇ PowerWeb M: 9.90 Euro/Month

    3 domains, 170 MB webspace, 170 POP3 mailboxes, unlimited data transfer, otherwise as before.

Web Hosting Offers (9)

- Strato Medien AG, continued:
  - ◇ Premium S: 19.90 Euro/Month.

    5 domains, 300 MB webspace (data mirrored on a RAID system, daily backups), 300 POP3 mailboxes, 3000 email aliases, 25 GB data transfer/month (0.04 Euro per additional MB), SSL encoding, own CGI programs, MS Frontpage 2002 Server Extensions, RealMedia Server, Live Picture Server (FlashPix), PHP 3, PHP 4, Phython, SSI, MySQL database, telnet access, anonymous FTP (incoming FTP), telephone support for 0.12 Euro/Min, Adobe GoLive 6.0, Adobe Photoshop Elements 2.0, Adobe LiveMotion 2.0.

  - ◇ Additional .de domain: 0.99 Euro/Month.

    Additional .com, .org, .net domain: 1.45 Euro/Month.
    .at/.ch domain: 3.99 Euro/Month plus 99 Euro one-time fee.

Web Hosting Offers (10)

- Loomes: [http://www.loomes.de]
- HostEurope: [http://www.hosteurope.de]
- Domainfactory: [http://www.domainfactory.de]
- Domainbox: [http://www.domainbox.de]
- FreeCity: [http://domains.freecity.de]

  The domain registration costs a one-time fee of 14.99 Euro. If one accepts advertisements (via pop-up window) for all accesses to the web pages under the domain, there is no further fee.

- Redirect to you: [http://www.r2u.de/]

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 (“webspace”)?
- 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 he 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 transfered 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 legitime 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.