

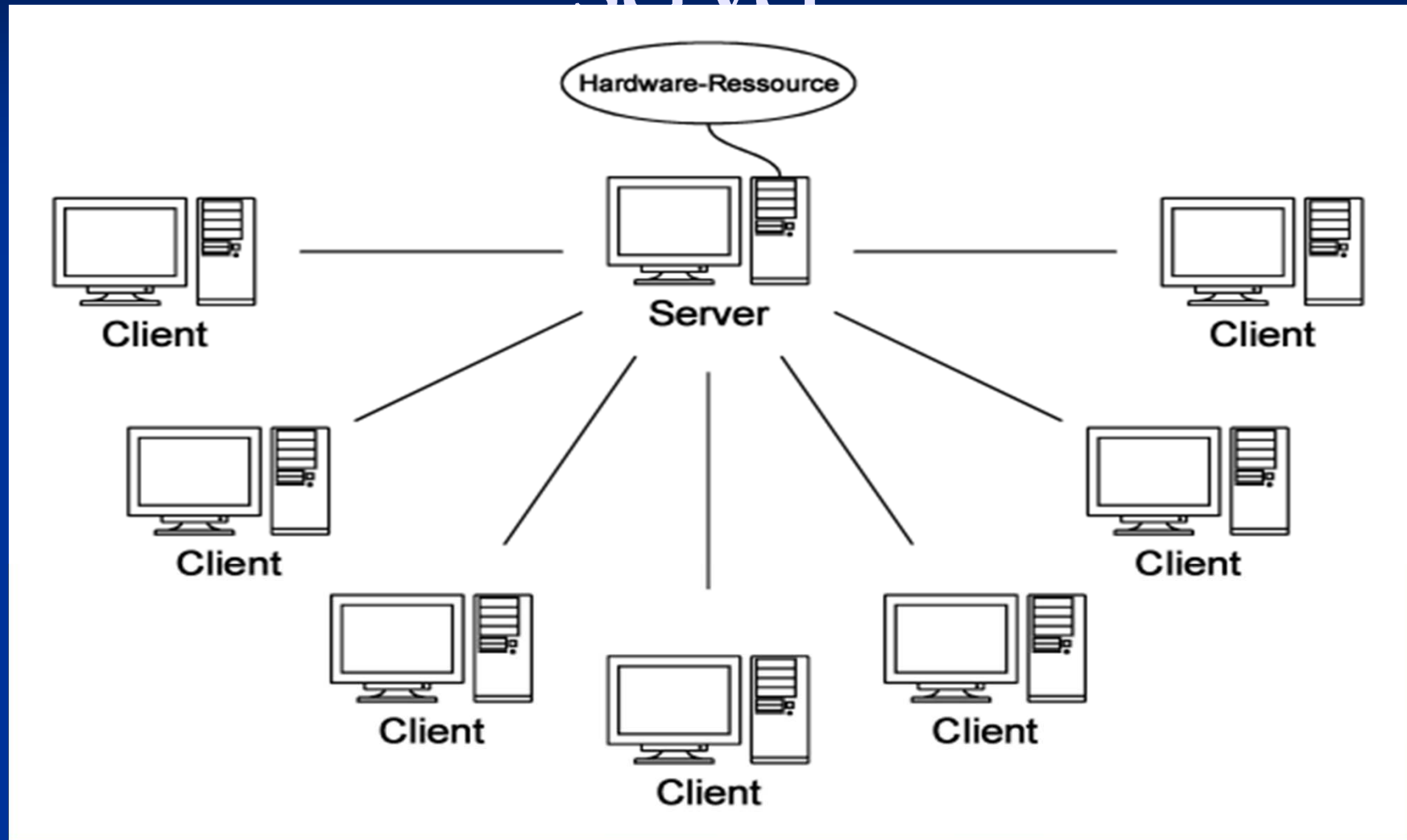
Internetrecherche



© 2011 Hans-Joachim Weller

http://www.amhochsitz.at/wp-content/uploads/2010/03/3951143570_20b4eccd3f.jpg

Die Seiten liegen auf einem Server



Domain

- ICANN oberster Verwalter, vergibt Netznamen
- TLD Top Level Domain, z. B. de
- NIC Network Information Center Registrar für Domains, z. B. DENIC = NIC.de
- IP Internet-Protokoll-Adresse
Z. B. 74.125.230.144 = google.de
- domain.de
domain = Name => Verweis auf IP-Adresse
- de = DENIC => Deutsche NIC
Adresse der aufgerufenen Seite
- sub.xxx.de = Subdomain einer Domain

Internetadresse

- **URL** Uniform Resource Locator
Adresse der aufgerufenen Seite
- **http:** hypertext transfer protocol
Protokoll der Kommunikation mit dem Server
gibt auch andere, z. B. ftp
- **www** world wide net
Protokoll für weltweite Vernetzung
- **: // . /**
Trennzeichen zwischen Zeichenketten
mit spezieller Bedeutung
- **Logbuch** (unsichtbar) Der Server notiert jeden Aufruf (IP,
Zeitstempel, aufgerufene Seite)

Adresselemente

- / Trennzeichen, danach kommt der Name eines Unterverzeichnisses / einer Datei
- Kommt danach nichts, liegt in diesem Verzeichnis eine Datei, die automatisch startet (index.html, default.aspx, etc.)
- ? Trennzeichen für Übergabeparameter
Protokoll für weltweite Vernetzung
- & Trennzeichen zu weiterem Parameter
- Parameter steuern die weitere Verarbeitung

Steuerelemente

- **Statische Seiten** z. B. .htm oder .html
=> haben immer den gleichen Inhalt
- **Dynamische Seiten** z. B. .php oder .aspx => bauen die Seiten beim Aufruf aktuell zusammen, daher sind die Seiten oft verschieden
- **Javascripte** ggf. extern als .js – Datei
=> Programmieren die Inhalte
- **Session-ID** wird für den Benutzer gesetzt, kann verfallen, danach geht die Seite nicht mehr
- **Links** Verweis auf andere Seite, weiterleitend oder neu aufgerufen, kann sich ändern oder scheitern (404)
- **Cookie** wird auf dem Client gespeichert. Informationen über den User (z. B. letzter Aufruf)
- **Flash Player**, etc. Für bestimmte Elemente müssen Zusatzprogramme installiert sein, es kann Versionsprobleme geben, Text kann Grafik sein

Seitenstruktur

<HTML>

<HEAD>

<TITLE> Titel der Seite </TITLE>

</HEAD>

<BODY>

<CENTER> <H1> Überschrift </H1> </CENTER>

normal, fette , <I> kursive </I>,
<U> unterstrichen </U>

 Zeilenumbruch

<IMG SRC=http://www.google.de/images/srpr/logo3w.png
HEIGHT="60" WIDTH="120">

 Link zu
Google

</BODY>

</HTML>

Suchelemente

- **And / Und / +** Elemente werden mit einander verknüpft => alle müssen wahr sein
- **OR** Elemente werden mit einander verknüpft => nur einer muss wahr sein
- **Not / -** => Element darf nicht zutreffen
- **„XXX YYY“** => Zeichenkette XXX YYY
- **STOPWörter** => übersprungene Wörter wie „und“ verwenden: „+und“
- Gross-/Kleinschreibung, Umlaute, Bindestriche werden ignoriert

erweiterte Suche

- Datum
- Sprache
- Dateityp
- In bestimmten Seiten
- Bestimmte Stelle in der Seite
- Bestimmte Länder
- Web, Bilder, Videos, News

Gefahren

- **Ranking** Reihenfolge nicht objektiv
- **Plagiate** Viele Inhalte sind entlehnt
- **Fehler** Keine redaktionelle Steuerung
- **Dubioses** kein Autor, dubiose TLD, kein Impressum, keine T&C, Weiterleitungen
- **Links** Gefahr von Viren, etc.
- **Tipp:** Whois verrät den Eigentümer

Auf Wiedersehen!