



max planck institut
informatik

Das Internet

Kurt Mehlhorn und Adrian Neumann
Max-Planck-Institut für Informatik

Vorlesung Ideen der Informatik

Folien von Kosta Panagiotou

Katzenvideos

- Katzenvideos sind ein integraler Bestandteil des Alltags

<https://www.youtube.com/watch?v=nCTey9W5YDc>

- Wie funktioniert das?



Überblick

- Datenübertragung
 - zwischen zwei Rechnern
 - zwischen Rechnern in einem Netzwerk
 - zwischen Netzen im Internet
- Aufbau von Webseiten
- Darstellung im Webbrowser

Datenübertragung

- Bits werden als Spannung am Kabel übertragen
- $+5V = 1$, $-5V = 0$
- ...Oder per WLAN
- ...Oder per Satellit
- ...Oder per Brieftaube
- Unterschiede müssen für den Benutzer unsichtbar sein!



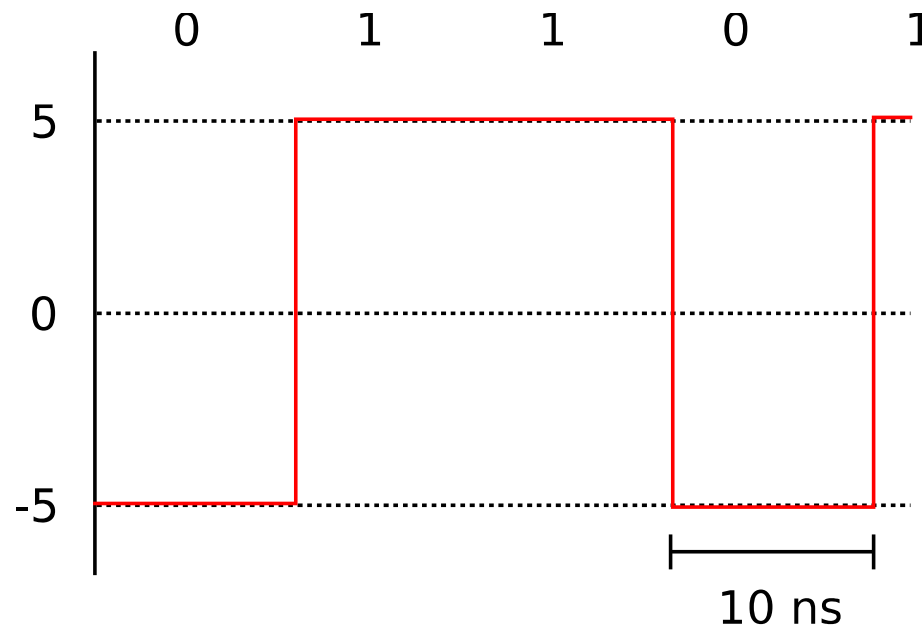
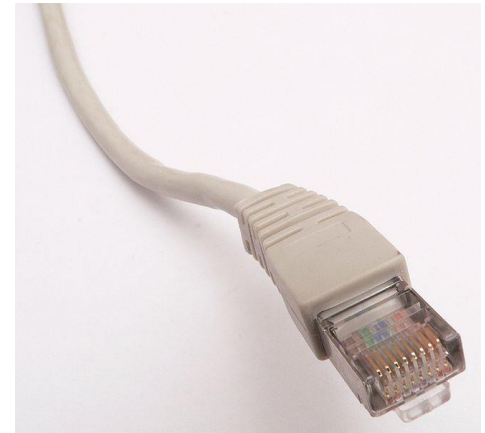
Schichten

- Link Layer
 - Abstrahiert von der Technik im lokalen Netz
 - Internet Layer
 - Verbindet das lokale Netz mit dem Netzanbieter, Transport ohne Garantien
 - Transport Layer
 - Fehlertolerante Datenübertragung
 - Data Layer
 - Kommunikationsprotokoll zwischen Browser und Server
- Text via Morsezeichen, Rauchzeichen, Fernschreiber, Brief,
 - die Post als Betrieb, Briefträger, Postsortieren, Transport mit Zug, Auto, Flugzeug,
 - die Postdienste, Einschreiben mit Rückantwort.
 - Versandhandel, Bestellungen, Vertrieb



Ethernet, eine populäres Netzwerk

- Kabelgebunden
- $+5V = 1$, $-5V = 0$
- 100-1000 Millionen Bits pro Sekunde



Probleme

- Uhren:
 - Wann messe ich die Spannung?
 - Welche Uhrenqualität braucht man?
 - 1000000 Einsen = 10^{-2} Sekunden 5V, nicht 10^{-2} Sekunden + 10 ns
- Störungen
 - Sollte das eine 1 sein, oder hat jemand den Fön angemacht?



Uhren

- Uhren mit Nanosekundenpräzision sind teuer;
- Lösung: Nie zu lange 1 oder 0 senden, z.B

ManchesterKodierung:

- Kodiere 0 als 01 und 1 als 10
- Also 0001101 als 01010110100110
- In der kodierten Folge, nie mehr als 2 gleiche Symbole hintereinander

Störungen

- Übertragungsfehler passieren ständig
 - 1 Fehler pro 10 Millionen Bits = 10 Fehler/s
- Meistens: Viele Bits hintereinander falsch
- Bits werden in Pakete zusammengefasst
- Jedes Paket bekommt eine Prüfsumme; siehe nächste Folie
- Bei Fehlern im Paket: Neuübertragung

Prüfsummen

- Einfachste Prüfsumme = Quersumme
- besser (Zahlendreher): gewichtet QS

Beispiel: Prüfziffer bei der ISBN-13

9 7 8 - 3 - 1 2 - 7 3 2 3 2 0 - ?

$$\begin{array}{cccccccccccc} \downarrow \times 1 & \downarrow \times 3 & \downarrow \times 1 & & \downarrow \times 3 & & \downarrow \times 1 & \downarrow \times 3 & \downarrow \times 1 & \downarrow \times 3 & \downarrow \times 1 & \downarrow \times 3 \\ 9 & 7 & 8 & + & 9 & + & 1 & 2 & + & 7 & 3 & 2 & 3 & 2 & 0 & = & 83 & \xrightarrow{\text{Abstand zum nächsthöheren Vielfachen von 10}} & 7 \end{array}$$

- Viele verschiedene Prüfsummen bekannt
Zum googeln:
 - Hamming codes
 - Cyclic redundancy checks
 - Reed-Solomon codes

MAC (media access control) Adressen

- Im Ethernet hört jeder alles auf der Leitung
- Jedes Gerät hat eine eindeutige MAC Adresse
- Datenpakete haben einen Addresspräfix
 - Wenn Adresse stimmt: Prozessor benachrichtigen
- Switches
 - Schauen in die Pakete
 - Leiten an das richtige Kabel



Internet Protocol

- Bietet Paket-Kommunikation *zwischen* Netzwerken
- Egal ob die Technik gleich ist (Ethernet vs. WLAN)
- Best Effort, Keine Garantien:
 - Pakete gehen verloren
 - Pakete kommen doppelt an
 - Reihenfolge kann sich ändern

IP Adressen

- Wie Telefonnummern für Computer
- 32 Bits für die Adresse
 - Vier Zahlen zwischen 0 und 255
 - Zum Beispiel *139.19.14.56 = MPI-INF*
 - Regionale Clustering
 - Hat man nicht von Geburt an (MAC-Adresse), sondern bekommt man zugewiesen
- Ungefähr 4 Milliarden mögliche Adressen
- Bald aufgebraucht: Umstieg auf 128 Bits



IP Routing

- Jeder Router (Verteiler) hat eine Tabelle
 - Pakete an 192.168.*.* gehen am besten an 192.168.0.01
- Ist Ziel in meinem Netz? Direkt an MAC
- Sonst in der Tabelle nachschlagen und weiterleiten

Routing Information Protocoll

- Router berechnen kontinuierlich kürzeste Pfade im Netz (kurz = wenige Hops)
- Wie in Vorlesung über kürzeste Wege
- Alle 30 Sekunden: Tabelle an alle Nachbarn weiterreichen
- Update für kürzere Routen

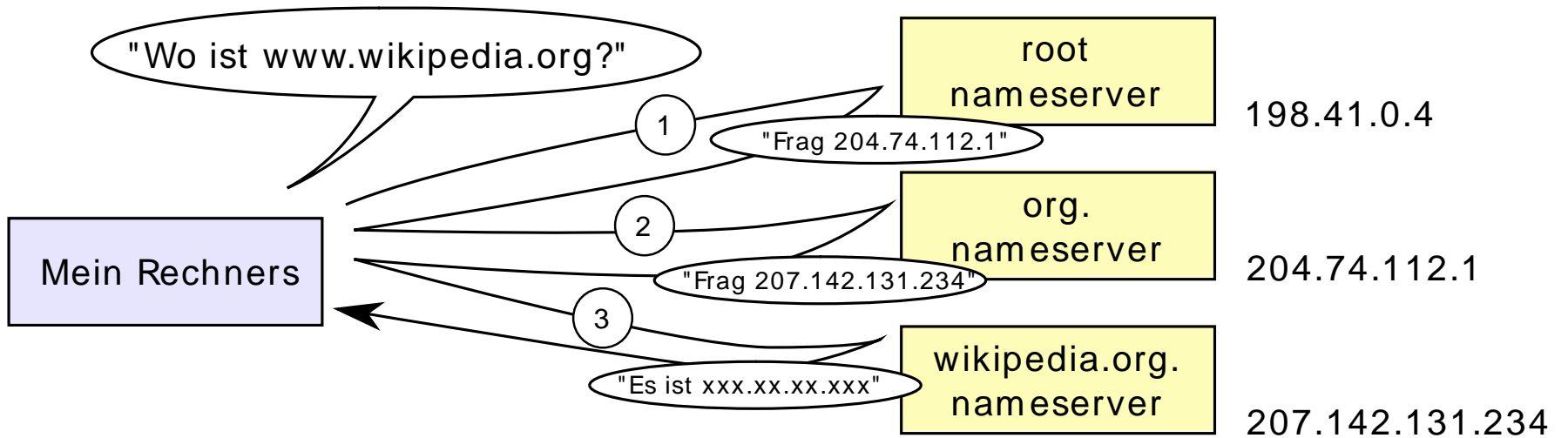
Transmission Control Protocol (TCP)

- Zuverlässige Datenübertragung zwischen Rechnern
 - Pakete nummerieren → Reihenfolge
 - Pakete mit Rückschein
 - Bleiben Bestätigungen aus → Neu senden



DNS

- Telefonbuch für IP Adressen
 - Übersetzt `www.google.de` in `173.194.35.151`
- “Nameserver” speichern Tabellen
 - Tabelle enthält entweder Paar (Name,IP)
 - Oder Verweis auf Nameserver (mit `.de` gehst du besser zur Telekom)
- Jeder Computer hat eine Liste mit Nameservern



Zwischenstand

- Ethernet um im lokalen Netzwerk zu reden
- IP um zwischen Netzwerken Pakete zu schicken
- TCP um zuverlässig über IP zu reden
- DNS um IP Adressen nachzuschlagen



Hypertext Transfer Protocol, HTTP

- HTTP ist ein Protokoll zur Übertragung von Daten auf der Anwendungsschicht über ein Rechnernetz.
- Es wird hauptsächlich eingesetzt, um Webseiten (Hypertext-Dokumente) aus dem World Wide Web (WWW) in einen Webbrowser zu laden.

Hypertext (HTML)

- “Sprache” in der Webseiten geschrieben sind
- Text mit Bildern, Verweisen, Buttons, klickbaren Objekten,

```
<html>
```

```
<body>
```

```
    <h1>Überschrift</h1>
```

```
    <p>Ein Absatz</p>
```

```
</body>
```

```
</html>
```

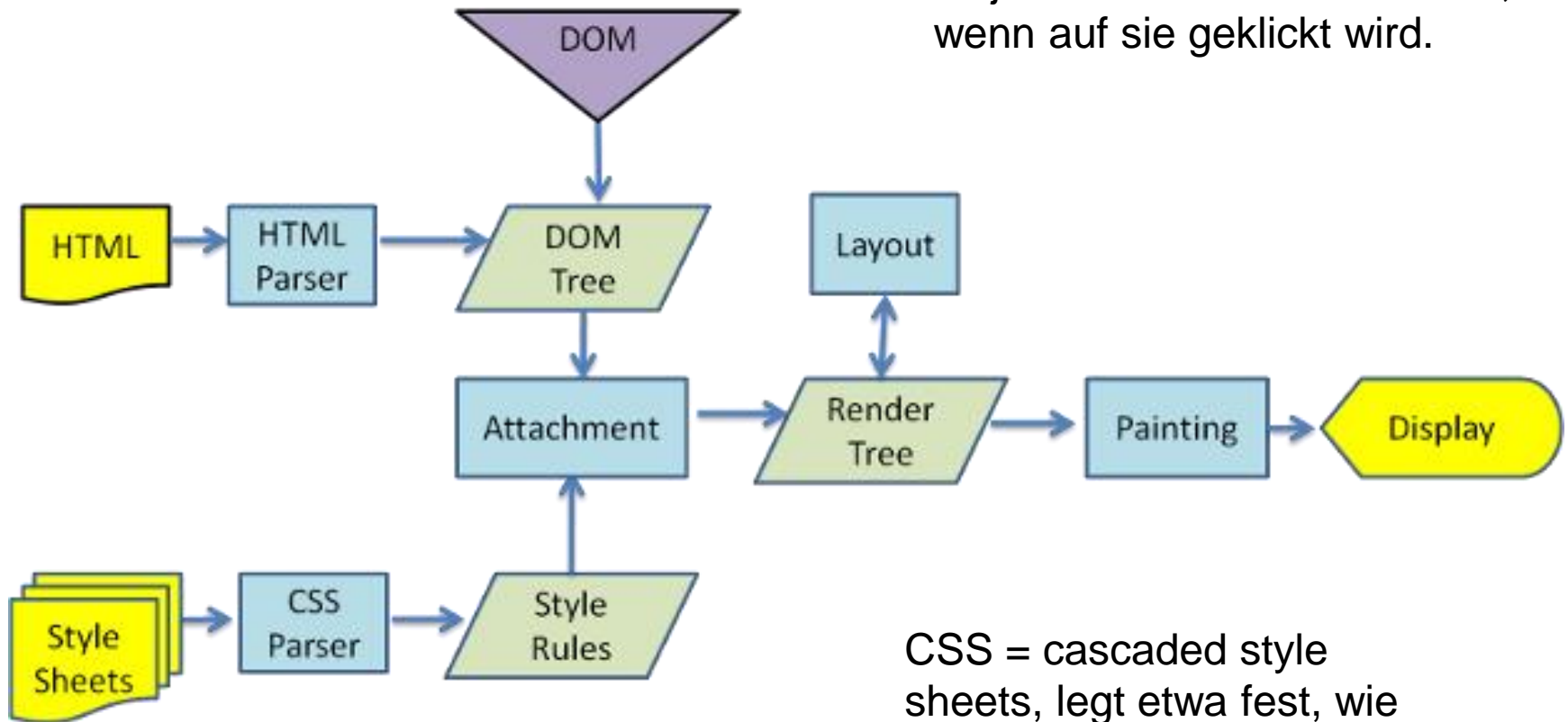
Überschrift

Ein Absatz

Darstellung im Browser

HTML ist reiner Text,
noch ohne Bedeutung.

DOM = domain object model,
Objekte wissen was zu tun ist,
wenn auf sie geklickt wird.



CSS = cascaded style
sheets, legt etwa fest, wie
Überschriften zu setzen
sind

Parsen

- Umwandeln von Buchstabensuppe zu hierarchischen Informationen und Objekten

```
<html>
```

```
<body>
```

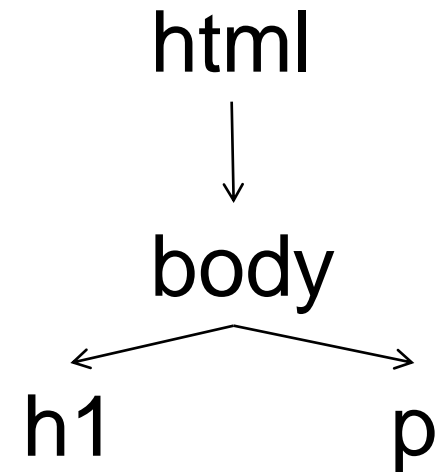
```
  <h1>Überschrift</h1>
```

```
  <p>Ein Absatz</p>
```

```
<a href="fotos/Kurt1980.jpg"> </a>
```

```
</body>
```

```
</html>
```



Zeichnen

- Das geparste Document wird “von oben nach unten” gezeichnet
- Weiter oben liegende Elemente definieren Stilkontexte
 - h1 verändert die Schriftdarstellung
 - p sorgt für Absatzvorläufe
 - a href sorgt dafür, dass man auf das Objekt klicken kann und was dann passiert.

Dynamische Elemente

- Mausbewegungen, Klicks etc. werden vom Betriebssystem verwaltet
- Browser wird über “Events” benachrichtigt
- Darstellung kann sich dynamisch ändern
 - Seite muss (effizient!) neu gezeichnet werden
- Klicken löst Aktionen aus
 - Zum Beispiel werden Videos abgespielt.

Videos

- Browser verstehen keine Videos
- Plugins
 - Werden vom Browser aufgerufen
 - Dürfen einen Teil der Webseite selbst zeichnen
- Für Videos heutzutage meistens 'Flash' von Adobe

Zusammenfassung

