

Funktionsweise des World Wide Web

Informatik - Seminar

Thema 5

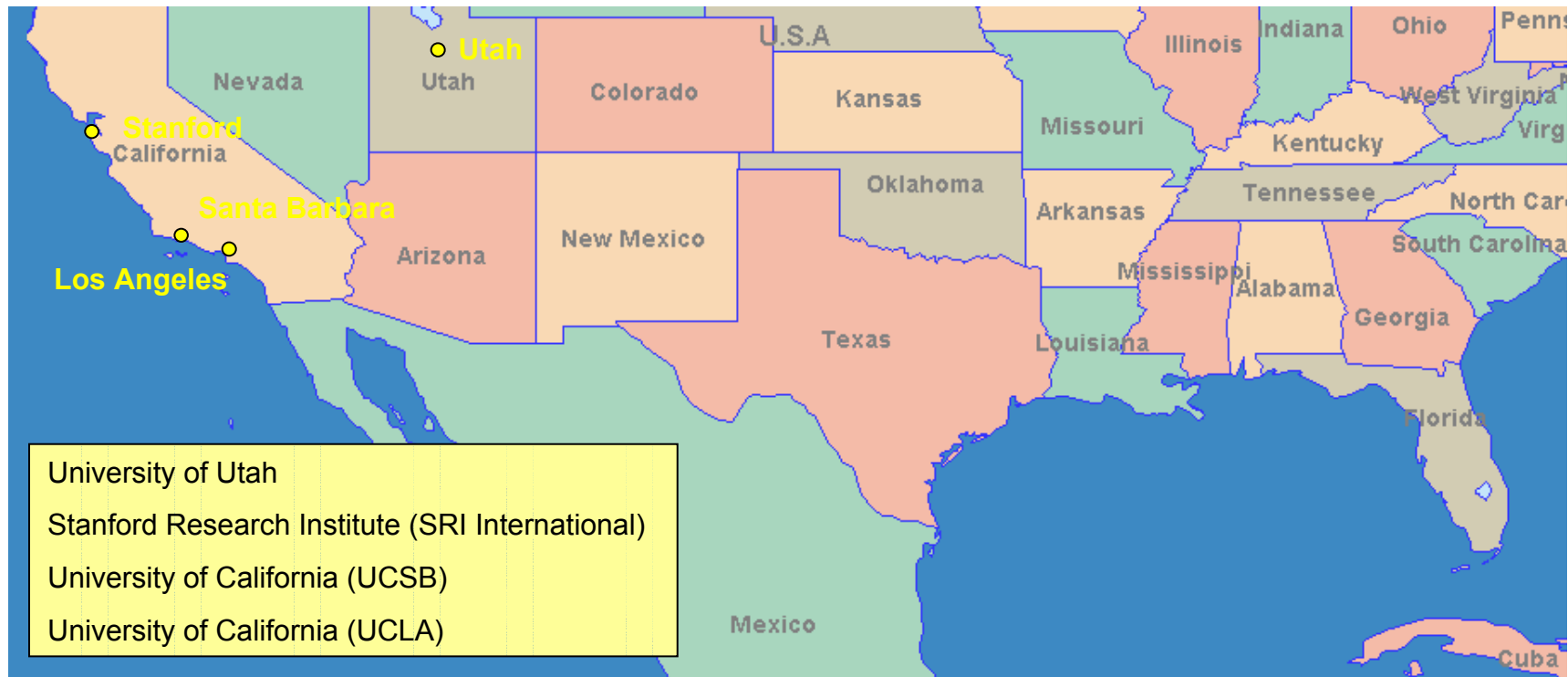
- **Entstehungsgeschichte**
- Physikalische Netzstruktur
- Kommunikationstechniken
- Domain Name Service
- Abschlussbetrachtung

- Das World Wide Web ist eine Ansammlung von Technologien, die ein verteiltes Hypermedia-Dokumentmodell auf Basis des Internets implementieren.
(Wilde 1999)
- Das Internet ist ein weltweiter Verbund vieler Millionen Rechner, die über das Netzwerkprotokoll TCP/IP miteinander vernetzt sind.
(Horn 1999)
- Das Web ist ein riesiges verteiltes System, das gemeinsame Ressourcennutzung durch offene Kommunikations- und Dokumentenstandards über verschiedene Dienste ermöglicht.

- 04.10.1957 Der erste Satellit Sputnik startet in den Weltraum
- 1958 Gründung der ARPA
(Advanced Research Projects Agency)
- 1965 Donald Davis entwirft die Paketvermittlung
- 1966 Entschluss zur Vernetzung der ARPA Großrechner
- 1969 Die ersten vier Hostrechner werden über IMPs
(Interface Message Prozessor) zum ARPANet
zusammengeschlossen
- 1969 - 1971 telnet, ftp, email
- 1972 Erste öffentliche Vorstellung vom ARPANet
Erweiterung beginnt

Entstehung vom Arpanet

- IMPs waren untereinander homogen
- ermöglichten den Aufbau des eigentlich heterogenen Netzwerkes



➤ Ab 1970 werden die IMPs durch normale Computer abgelöst

- 1983 Trennung von Arpanet in militärischen Teil MILNet und zivilen Teil (weiterhin) ARPANet
- 1983 NCP wird durch TCP/IP abgelöst
- 1986 Gründung der NSF (National Science Foundation) und des NSFNet
Ziel: Netzzugang für alle amerikanischen Universitäten
- 1989 Wird ARPANet vom DOD (Department of Defense) aufgelöst
- 1990 Geburtsstunde vom Web
CERN Mitarbeiter Tim Berners-Lee
Idee: Dokumente mit tagged language, Hyperlinks, http
Entwicklung des ersten grafischen Webbrowsers



1991 Gründung der ISOC (Internet Society)
zwecks Weiterentwicklung, heute oberste
Dachorganisation



Wichtigste Unterorganisation ist das IAB
Definition der „Internet-Standards“
Veröffentlichung und Verwaltung der RFCs

RFCs (Requests for Comments) beschreiben die
aktuelle Internetarchitektur.

Herstellerneutral, offen, lizenzfrei
Jeder darf sich beteiligen und Vorschläge machen

Zur Zeit 3776 RFCs (Stand April 2004)

- Entstehungsgeschichte
- **Physikalische Netzstruktur**
- Kommunikationstechniken
- Domain Name Service
- Abschlussbetrachtung




Zugang über ISP (Internet Service Provider)
Ablauf: Einwahl und Authentifizierung

Telefonnetz Modem, ISDN, DSL-Varianten

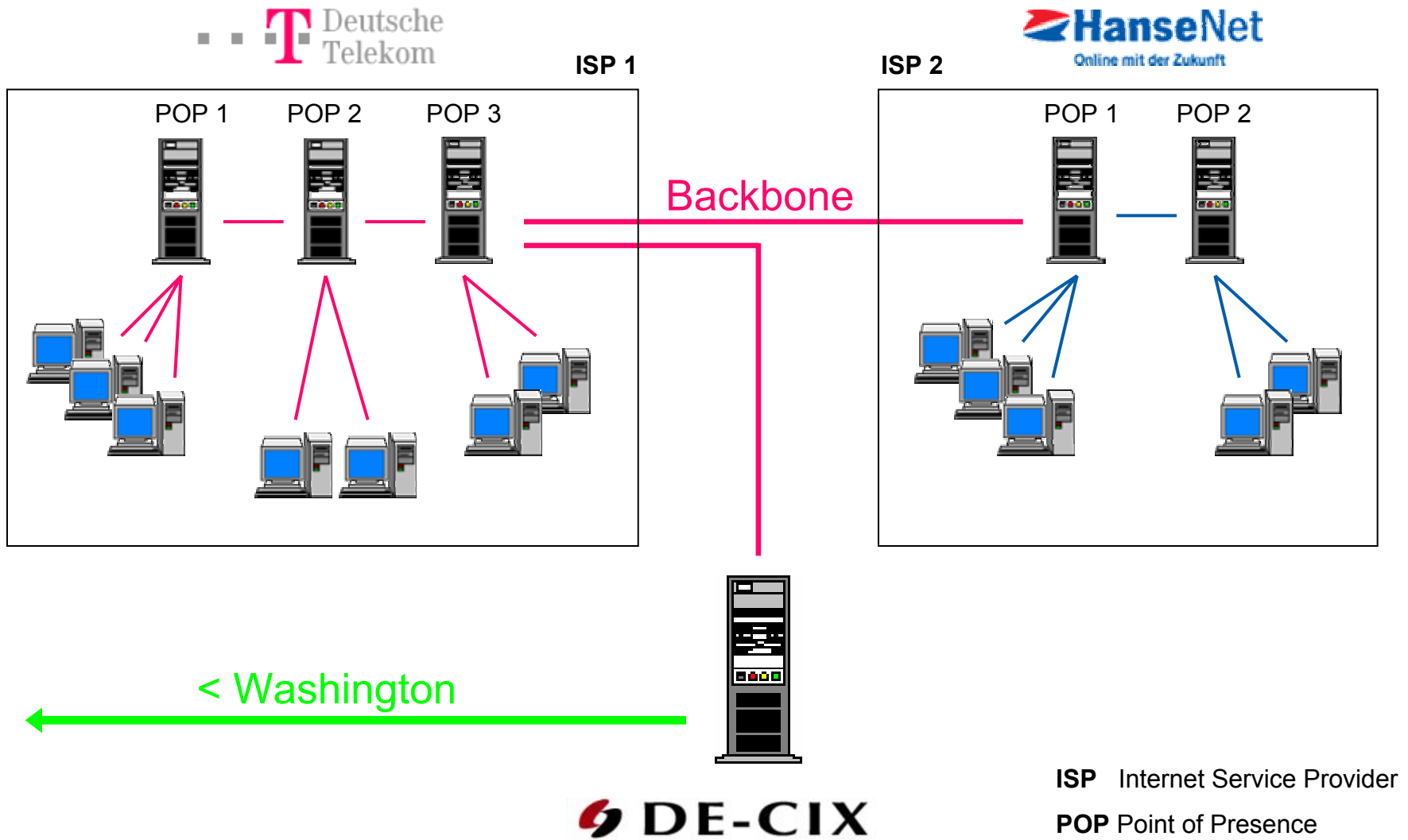
Mobilnetz GSM, GPRS, UMTS

Stromnetz PLC (Powerline Communications)

TV-Netz TV-Kabelmodem, noch keine Standardisierung erfolgt

- Vergabe von IP-Adressen erfolgt hierarchisch und blockweise
- IANA (Intermodal Association of North America) vergibt und verwaltet den gesamten IP-Adressraum und vergibt IP-Adressblöcke an RIRs (Regional Internet Registries)
Europa, Nordafrika  www.ripe.net
- Provider können bei RIRs IP-Adressblöcke beantragen
Autonome Verwaltung der IPs, Weitergabe möglich
- Jeder Rechner erhält für die Dauer seiner Internet-Verbindung eine weltweit einmalige IP-Adresse.
- IP = 32-bit Adresse, in DDN (Dotted Decimal Notation) Dezimale Darstellung, in der die 4 Bytewerte durch Punkte getrennt werden 213.39.232.194 (www.fh-wedel.de)

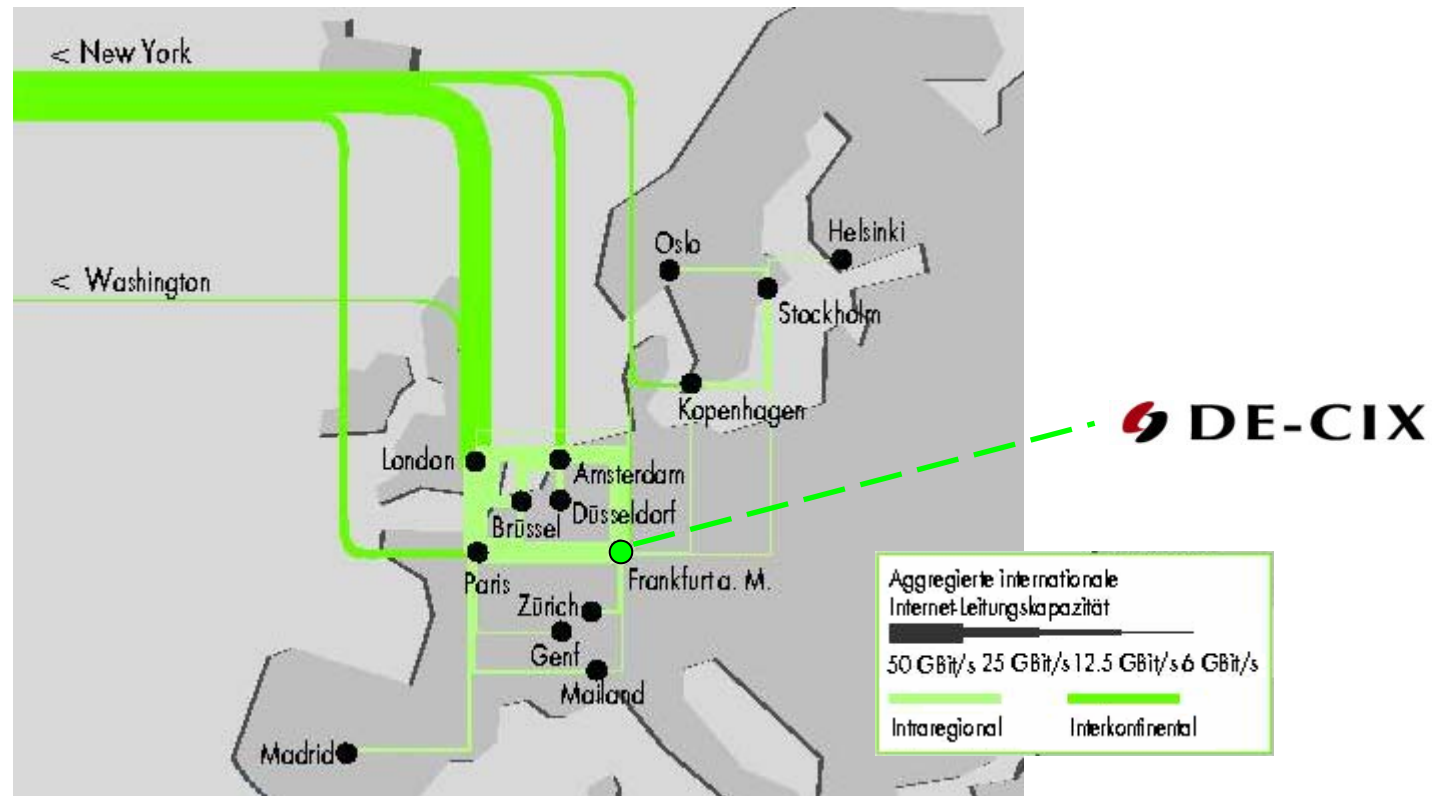
Entstehung vom Web



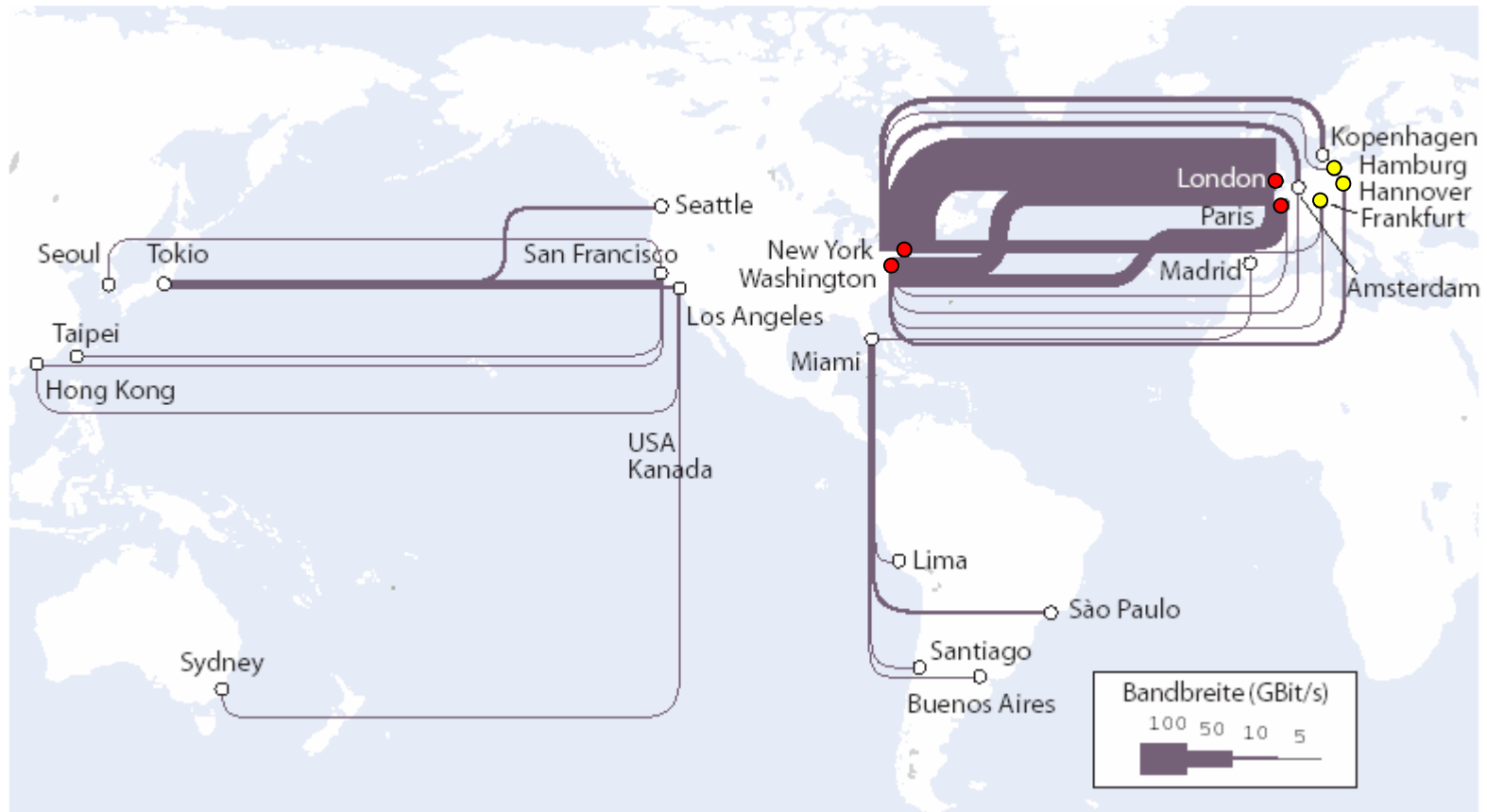
Backbone (Rückrat)

- Hochgeschwindigkeitsnetz, das die Netze der ISPs (Internet Service Provider) miteinander verbindet.
 - Backbone-Betreiber werden auch IP-Carrier genannt
-
- Nahbereich:
Deutsches Forschungsnetz (DFN)
Deutsche Telekom
 - International:
British Telecom, Telecom, France Telecom

- Größter deutscher Austauschpunkt ist das DE-CIX in Frankfurt



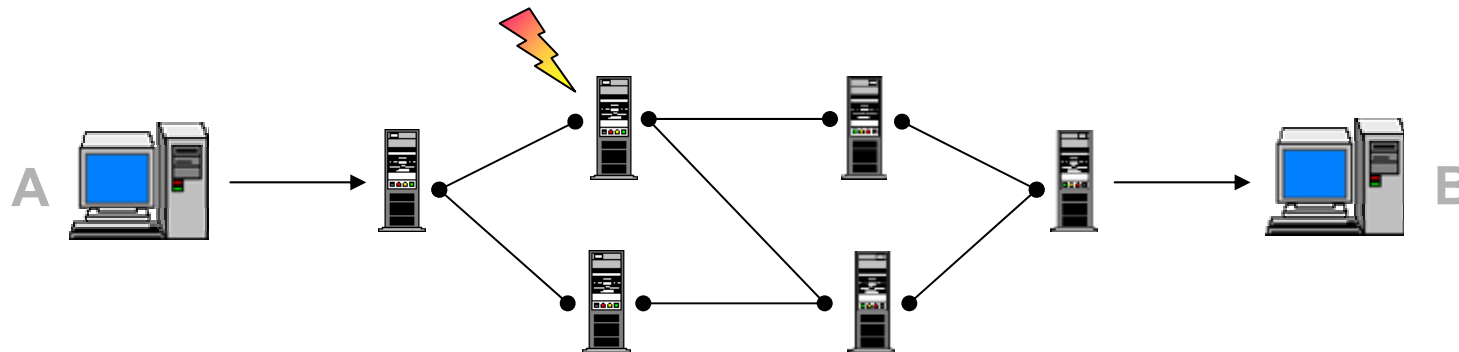
- 90% aller deutschen IP-Pakete gehen durch den DE-CIX Switch



- Konzentration auf nur 4 Knotenpunkte

- Entstehungsgeschichte
- Physikalische Netzstruktur
- **Kommunikationstechniken**
- Domain Name Service
- Abschlussbetrachtung

- Daten werden in einzelne Datenpakete zerlegt (bis max. 1500 Bytes)
- Datenpaket enthält Header mit Herkunftsadresse und Zieladresse sowie weiteren Metainformationen
- Router transportieren Pakete, keine festgelegten Übertragungsrouten



- Pakete können im Netzwerk verschiedene Wege durchlaufen
- Pakete können in unterschiedlicher Reihenfolge ankommen

- Pakete können beschädigt werden oder verschwinden
- Pakete werden beim Empfänger in richtiger Reihenfolge zusammengesetzt und bei Fehler neu angefordert

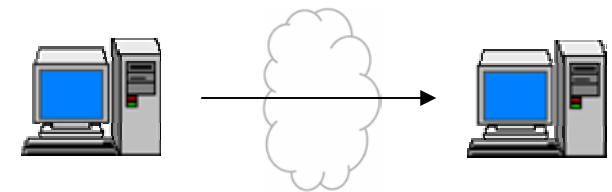
Vorteile

- mehrere Rechner teilen sich eine Verbindung
- Paketvermittlung ermöglicht allen Teilnehmern gleichberechtigten Zugang
- Pakete sind klein, Wartezeit ist gering
- Aufteilung der Daten in kleine Blöcke erleichtert die Entdeckung von Übertragungsfehlern

Protokolle regeln den Datenaustausch zwischen den verschiedenen kommunizierenden Systemen

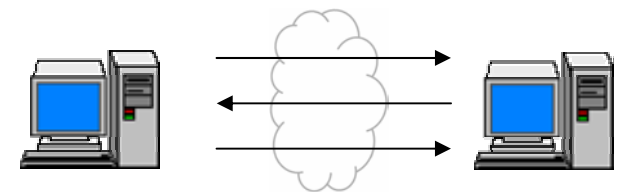
Verbindungslose Protokolle

- Pakete werden schnellstmöglichst übertragen
- Kein vorhergehender Verbindungsaufbau
- Keine Bestätigung für den Empfang
- Fehler (ICMP) DNS, time (UDP)



Verbindungsorientierte Protokolle

- Verbindungsaufbau
- Datenübertragung
- Verbindungsabbau



Exterior Gateway Protocol (EGP)

- Abfrage der Nachbarn alle 120-480 Sekunden
- Bei Änderung Austausch der kompletten Routing-Tabelle

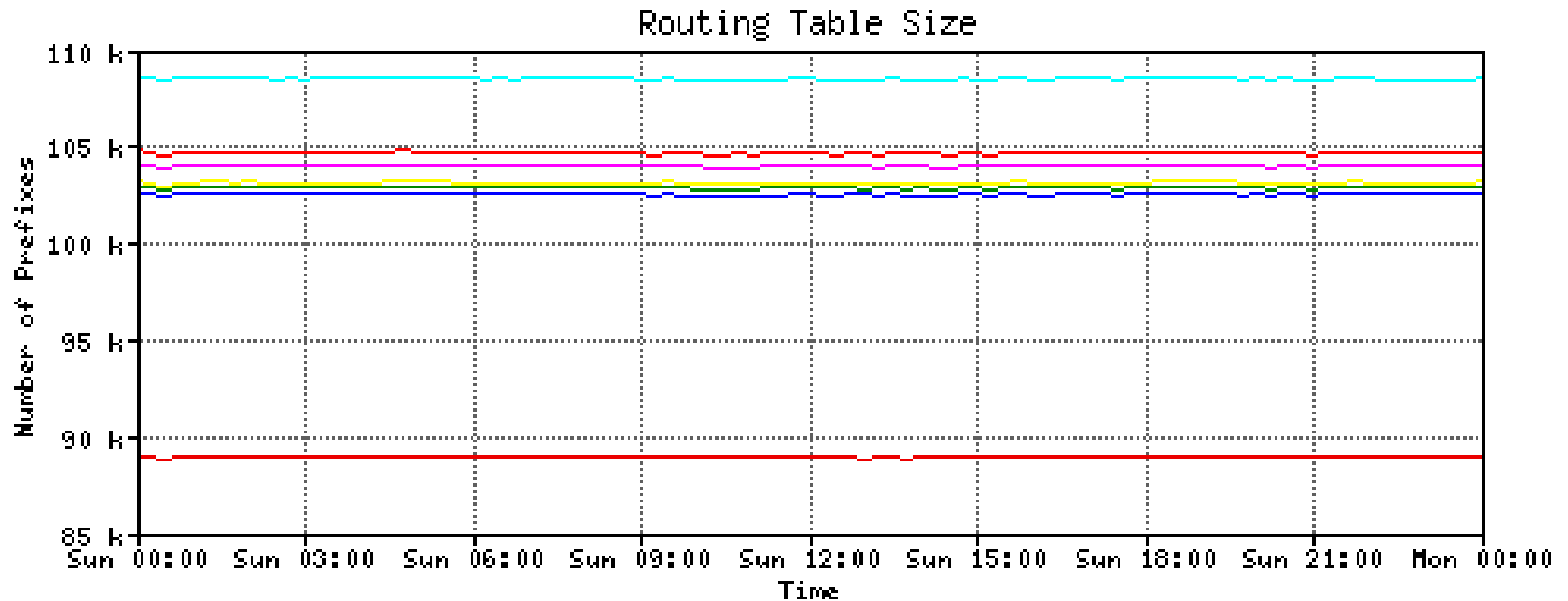
Border Gateway Protocol (BGP)

- Bei Änderung, Verschickung der Änderung an Nachbarn

Aufbau Routing-Tabelle

- Liste aller bekannten Netze und wie sie erreicht werden können
- Adressen, die über diese Router erreicht werden können
- Cost Metric (Bewertung der einzelnen Pfade: KBit/s, Auslastung)

➤ Gegenwärtig findet ein Übergang von EGP zu BGP statt



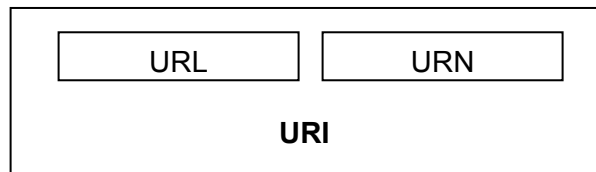
- Differenzen bis zu 20 Kb
- Bis zu 5% des Netzes zeitweilig nicht erreichbar

- HTTP ist das Übertragungsprotokoll für Hypertext-Dokumente im Web
- alle Adressen aller Websites beginnen mit `http://`
- Baut auf dem Daten-Transportprotokoll TCP/IP auf
- http besitzt Anforderungsbefehle wie `GET`
- Daten werden zeichenorientiert übertragen
- Zusätzlich Übertragungsparameter und Rückgabewerte
- keine dauerhafte Verbindung, für jede Anfrage neue Verbindung
zustandsloses Protokoll

URL (Uniform Resource Locator)


- Bisher ist nur der Uniform Resource Locator im Einsatz, die weiteren Formen sind noch in der Diskussionsphase. (Unger 2004)
- URI und URL bei typischen Adressen faktisch das Gleiche (selfhtml)

```
schema ":" "://" [user [":"password] "@"] host [":" port] "/" url-path
```



URN (Uniform Resource Name)

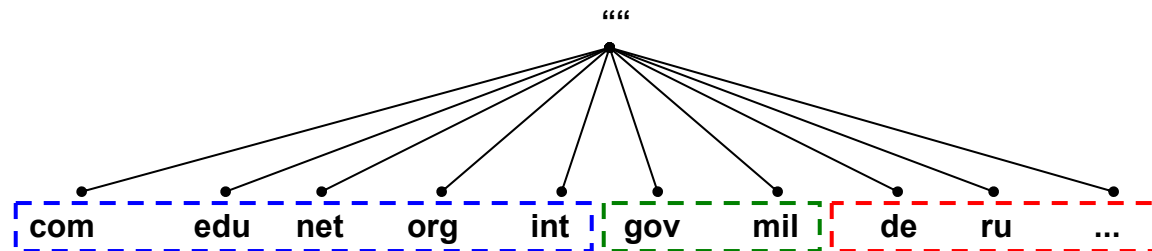
- Dauerhafte Adressierung wie bei ISBN
- Zum Auflösen UNS (Universal Naming Service) benötigt
- Ähnlich wie DNS zur Auflösen von Domainnamen in IPs (Meinel/Sack)




- Standardisierte Auszeichnungssprache, die auf SGML basiert.
SGML (Standard Generalized Markup Language) schon 1986
- Aktuelle Version: HTML 4.01 (91 Tags, 120 Attribute)
- Einfachheit, Erstellen mit beliebigem Editor, Darstellen im Browser
- plattformunabhängig, herstellerneutral
- Trennung von Daten und Layout
dafür Formatsprache CSS
- Standardisierung und Weiterentwicklung durch das
World Wide Web Consortium  <http://www.w3.org/>

- Entstehungsgeschichte
- Physikalische Netzstruktur
- Kommunikationstechniken
- **Domain Name Service**
- Abschlussbetrachtung

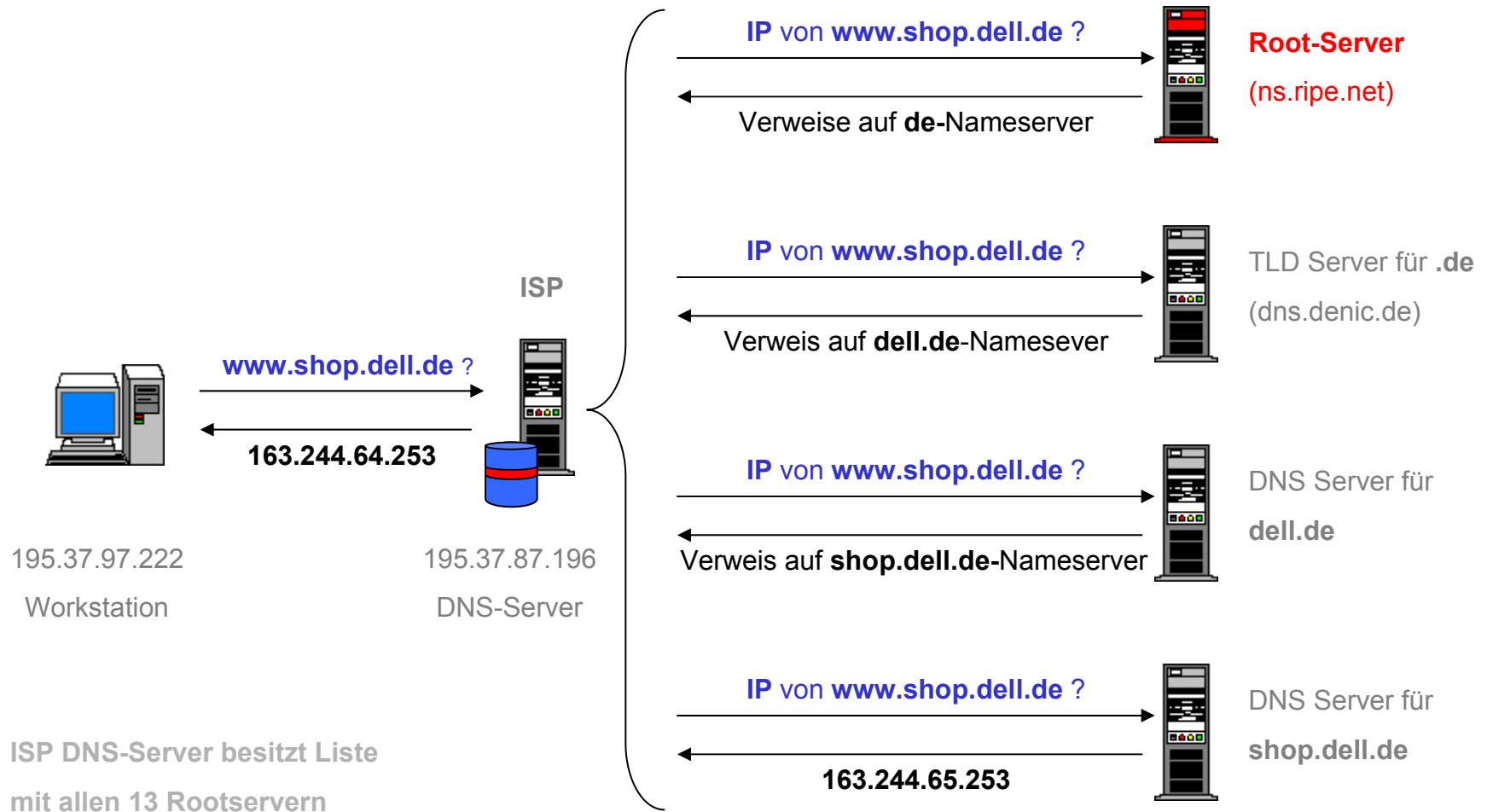
Organisation des DNS Name Space

- Hierarchischer Aufbau
- Aufteilung in drei Domainbereiche:



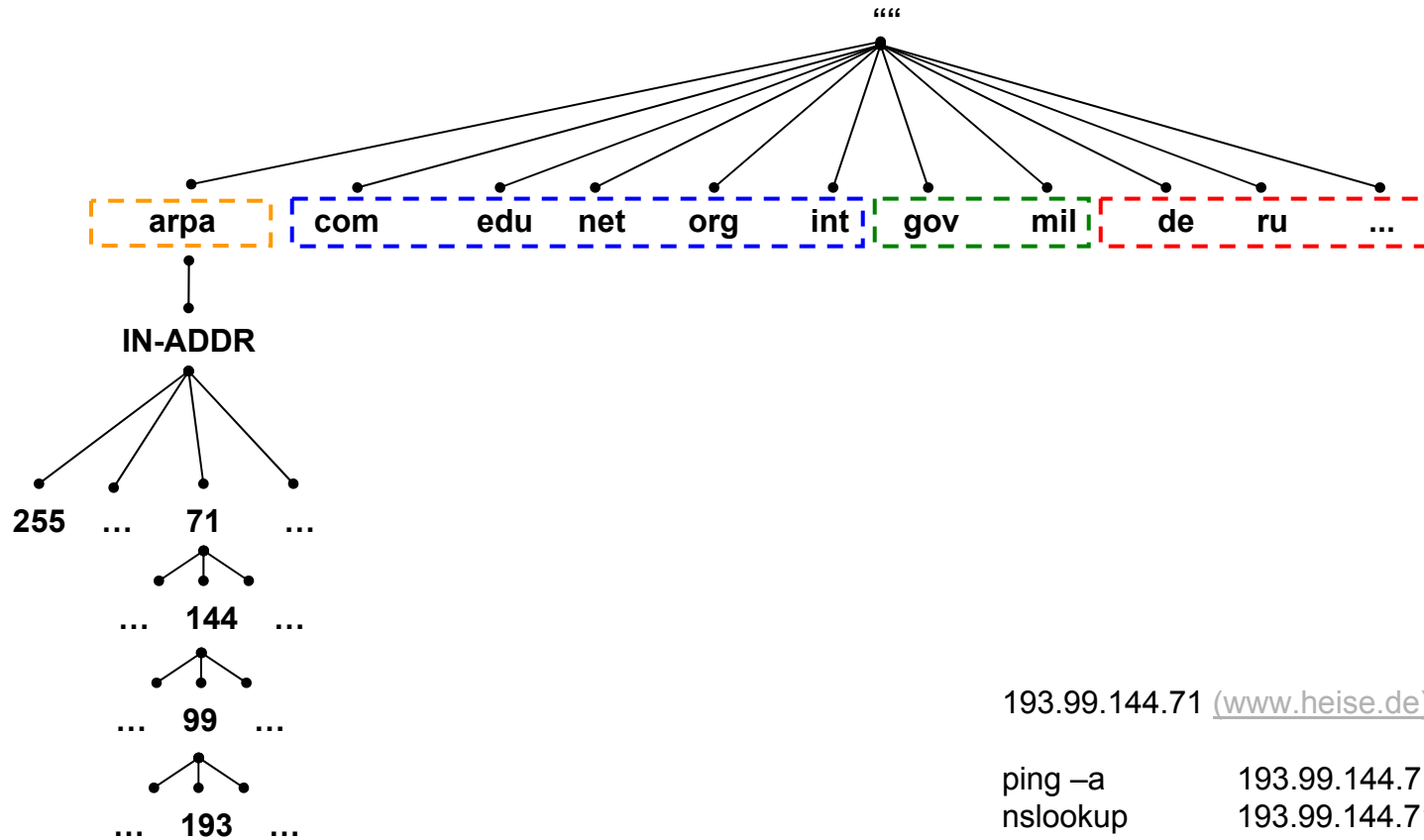
-  Generische weltweite Domains
-  Generische Domains der USA
-  Landesspezifische Domains

Vorwärtsauflösung (URL → IP)



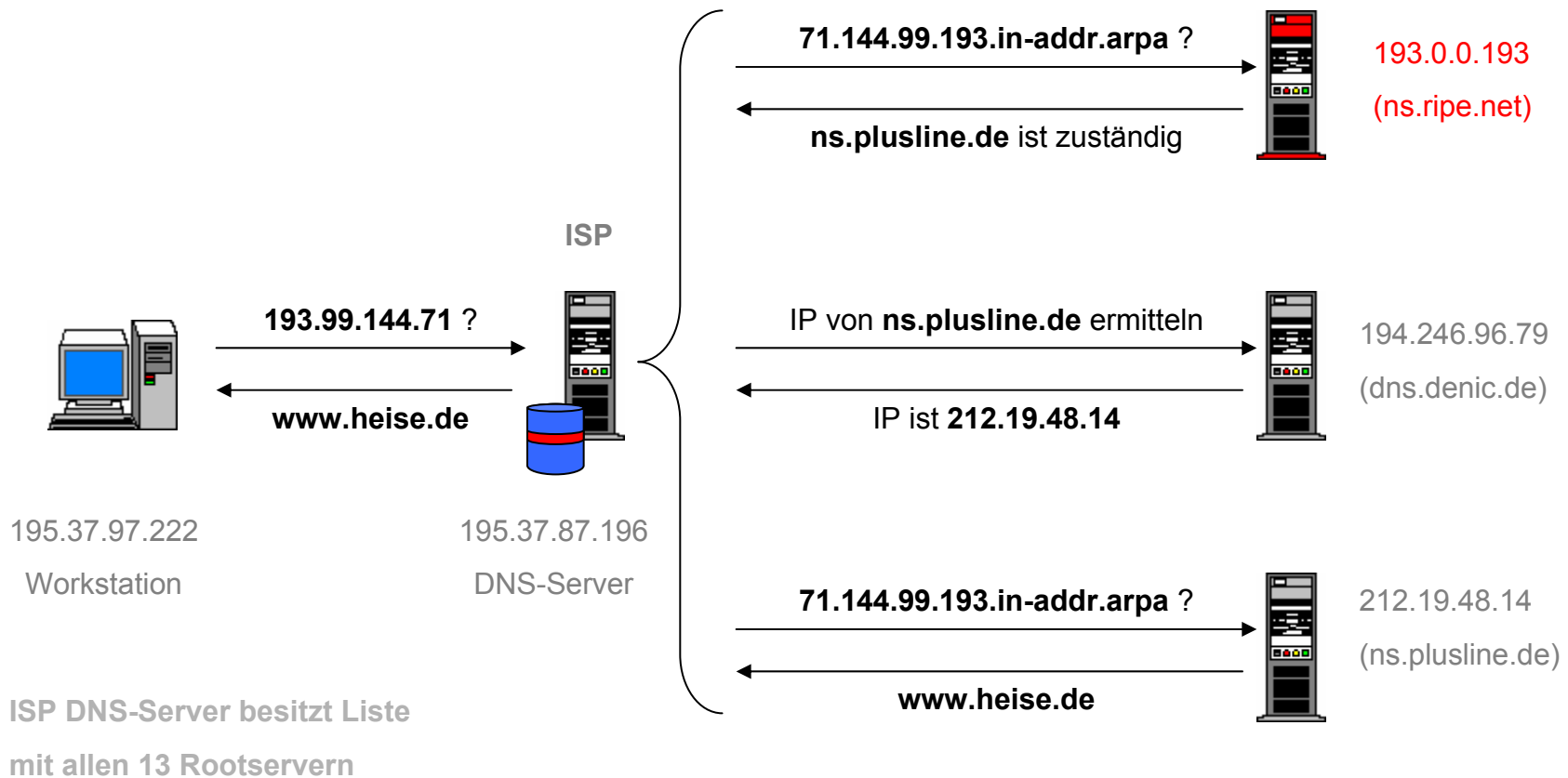
Erweiterung des DNS Name Spaces

- Rückwärtsauflösung (IP → URL)

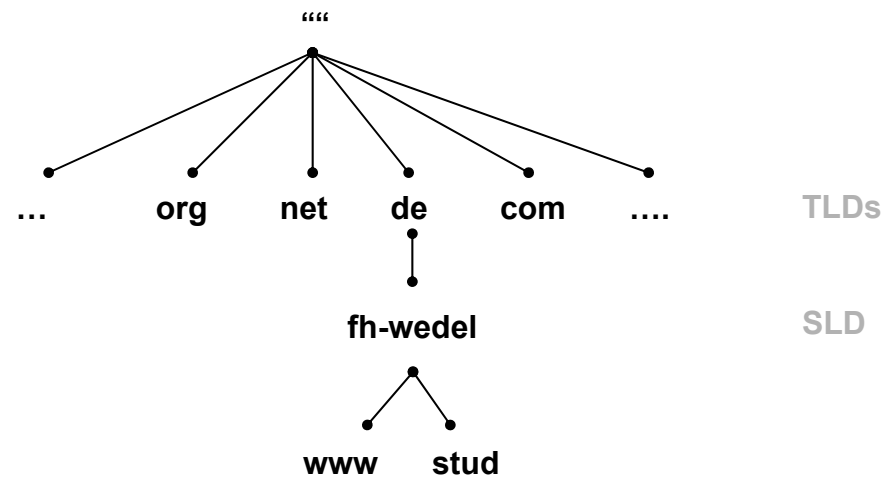


Rückwärtsauflösung

- Rückwärtsauflösung (IP → URL)



- Registriert wird nur die Domäne fh-wedel.de (ohne www davor)
standardmäßig ist diese über www.fh-wedel.de erreichbar



- Möglichkeiten mit eigenem DNS-Server:

<http://ftd.de> <http://stud.fh-wedel.de> <http://www2.din.de>

13 DNS Root-Server

- Weltweit existieren nur 13 Root Nameserver



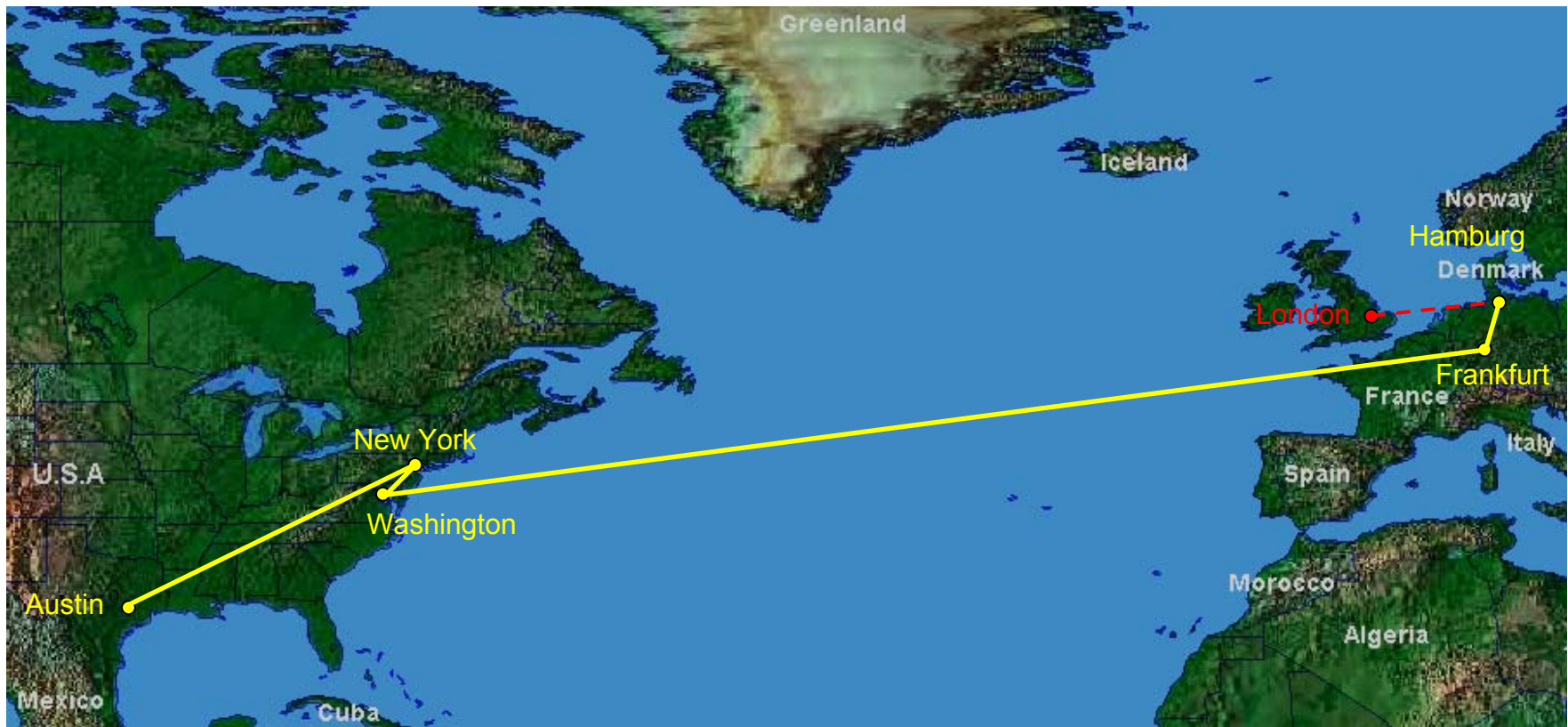
- 10 davon stehen in den USA

13 DNS Root-Sever

Name	Organization	City, State/Province	Country
A	Network Solutions, Inc	Herndon, VA	USA
B	Information Sciences Institute, University of Southern California	Marina Del Rey, CA	USA
C	PSINet	Herndon, VA	USA
D	University of Maryland	College Park, MD	USA
E	National Aeronautics and Space Administration	Mountain View, CA	USA
F	Internet Software Consortium	Palo Alto, CA	USA
G	Defense Information Systems Agency	Vienna, VA	USA
H	Army Research Laboratory	Aberdeen, MD	USA
I	NORDUNet	Stockholm	Sweden
J	(TBD)	Herndon, VA	USA
K	RIPE-NCC	London	UK
L	(TBD)	Marina Del Rey, CA	USA
M	WIDE	Tokyo	Japan

Konkretes Fallbeispiel

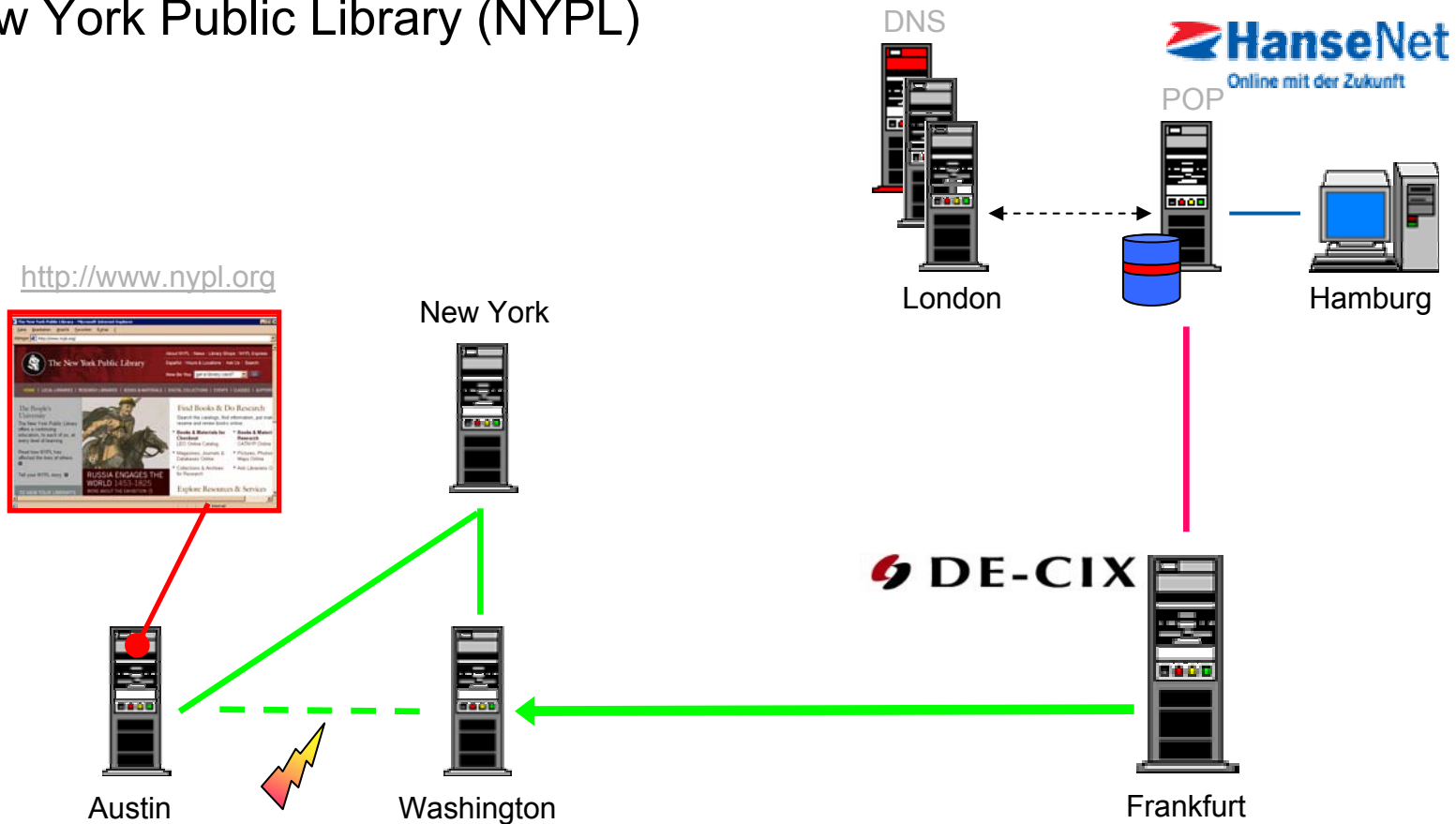
Verbindungsaufbau zu <http://www.nypl.org>



Konkretes Fallbeispiel

Verbindungsaufbau zu <http://www.nypl.org>

New York Public Library (NYPL)



- Entstehungsgeschichte
- Physikalische Netzstruktur
- Kommunikationstechniken
- Domain Name Service
- **Abschlussbetrachtung**

- Entwicklung neuer Dienste und Nutzungsmöglichkeiten
- World Wide Web wird weiter wachsen
- Konzentrationspunkte werden vorerst bestehen bleiben
- Top Level Domains werden erweitert
- IPs werden ebenfalls knapp (IPv6)

**Vielen Dank für Ihre
Aufmerksamkeit**