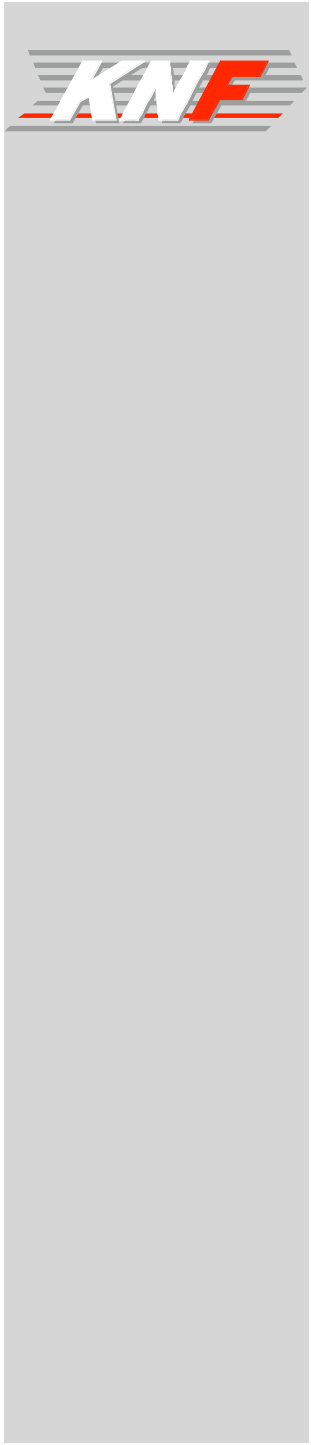


DNS

Einblicke in einen wenig bekannten aber viel genutzten Dienst im Internet

Jörg Kinzebach
<joerg@duck.franken.de>

DNS

Domain Name System

- RFC 1034 „Domain names - concepts and facilities“
- RFC 1035 „Domain names - implementation and specification“
- RFC 799 „Internet Name Domains“
- RFC 3845 „DNS Security (DNSSEC)“
- RFC 2136 „Dynamic Updates in the Domain Name System“ (DNS UPDATE
- RFC 1982 „Serial Number Arithmetic“
- RFC 1995 „Incremental Zone Trasnfer in DNS“
- RFC 1996 „A Mechanism for Prompt Notification of Zone Changes (DNS NOTIFY)
- RFC 1912 „Common DNS Operational and Configuration Errors“
- RFC 1591 „Domain Name System Structure and Delegation“

DNS - Namensauflösung

- Forward (IN A)

  www.franken.de ?

  ⇒ 193.175.24.24
- Reverse (IN PTR)

  193.175.24.33 ?

  ⇒ dns.franken.de
- Andere
  - MX für franken.de
    - franken.de. IN MX 10 laurel.franken.de.
    - franken.de. IN MX 20 priscilla.franken.de.
  - TXT für franken.de
    - franken.de. IN TXT "v=spf1 mx ip4:193.175.24.0/24 ip4:193.174.2.0/24 ip4:195.37.132.0/24 ip4:193.174.159.0/24 -all"

DNS - Historie

- Verwaltung als hosts.txt (/etc/hosts)
- Verbreitung via FTP

- Traffic stieg durch Anzahl der Hosts
- Lokale Netzwerke wollten lokal verwaltet werden
- Neue und kompliziertere Anwendungen erforderten besseres System zur Namensauflösung

DNS - Historie

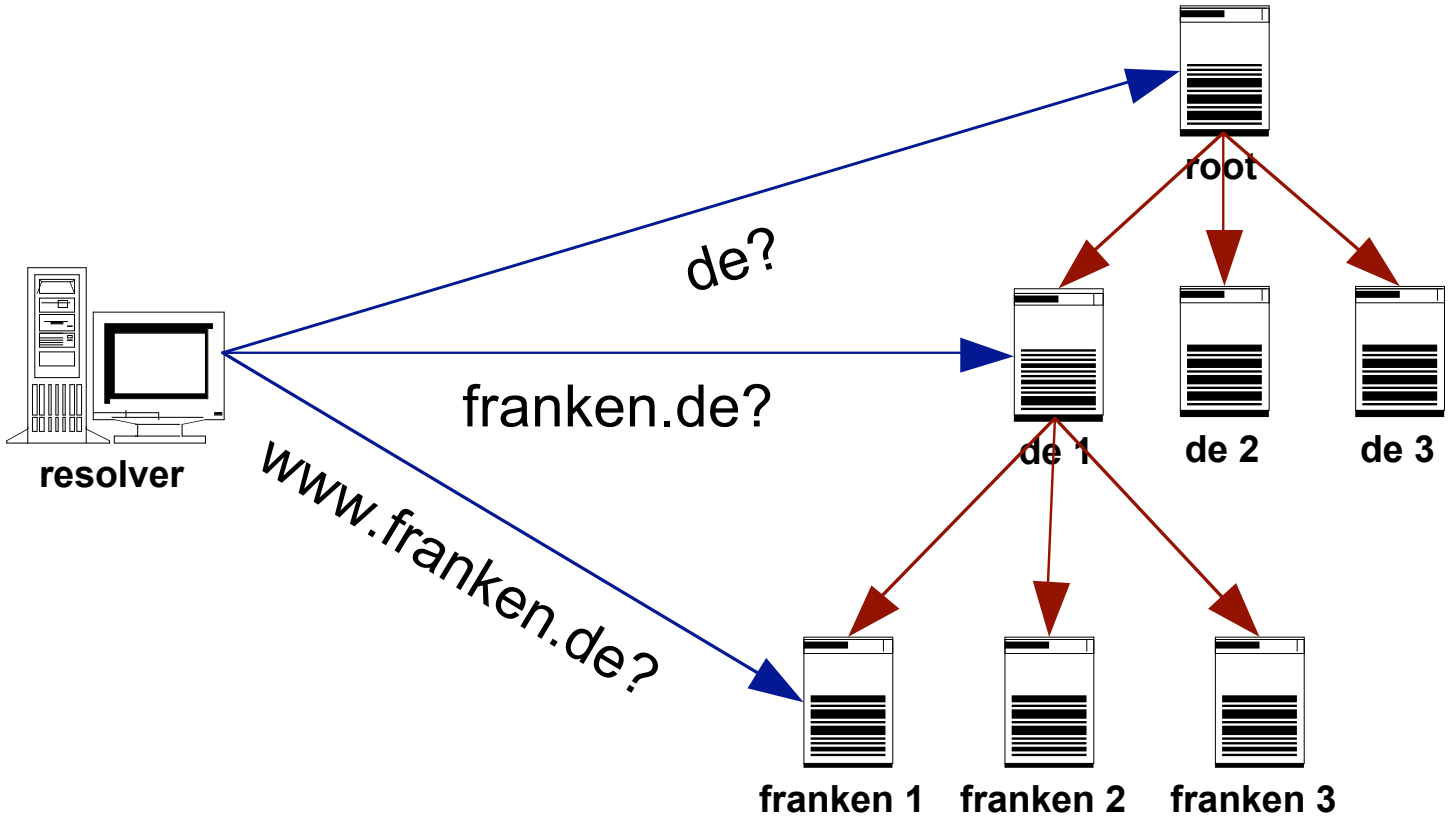
- Erdacht und spezifiziert 1983 durch Paul Mockapetris
- Erste Implementation genannt „Jeeves“ in 1983 auf dem TOPS-20 Betriebssystem
- BIND entstand an der UCB in einem Projekt für eine UNIX Implementierung für VAX Rechner
- Mitte der 90‘er wurde ISC gegründet um die Weiterentwicklung zu fördern

DNS - Workings

- Client fragt bei Resolver
- Resolver „hangelt“ sich durch DNS um eine Antwort zu erhalten
  - Baum-Struktur
  - Redundanz von Server
- Client erhält Antwort von Resolver

  - Timeouts
  - Caching
  - Lame Server

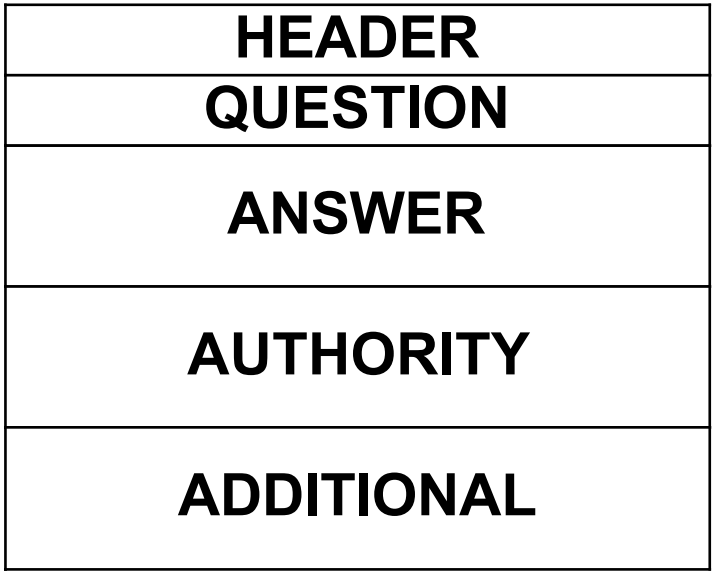
DNS - Baumstruktur



DNS-Pakete

- UDP
  - Maximal 512 Bytes.
  - Die meisten Anfragen sind UDP
- TCP
  - Für Anfragen und Antworten, die die Längenbeschränkungen von UDP Paketen überschreiten
  - Alle Zone-Transfers erfolgen via TCP

Packet-Format



- Header: Typ der Anfrage/Antwort und Anzahl der Einträge in den Sektionen
- Question: die Anfrage an den Nameserver
- Answer: RR mit Antworten zur Anfrage
- Authority: Verweise auf Server, die definitive Antworten liefern können
- Additional liefert hilfreiche Zusatzinformationen

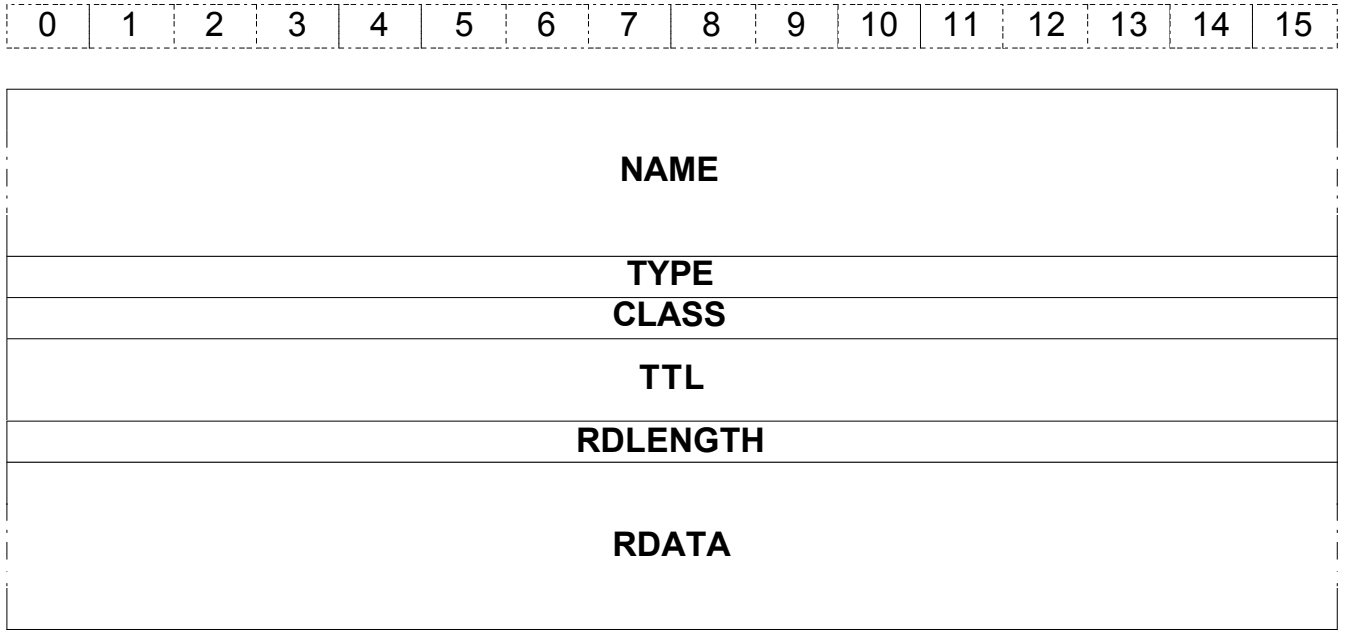
Packet-Header

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

ID							
QR	**OPCODE**	**AA**	**TC**	**RD**	**RA**	**Z**	**RCODE**
QCOUNT							
ANCOUNT							
NSCOUNT							
ARCOUNT							

- QR: Question/Response 0=request 1=response
- OPCODE: 0=standard query (QUERY),1=inverse query (IQUERY), 2=server status (STATUS), 3-15=reserved
- AA: Authorative Answer
- TC: Truncated
- RD: Recursion Desired
- RA: Recursion Available
- Z: Reserved
- RCODE: 0=no error, 1=format error, 2=server failure, 3=name error (bei AA=1: does not exist), 4=not implemented, 5=refused, 6-15: reserved

Resource Record (RR)



- NAME: „domain name“
- TYPE: spezifiert den Typ des RR und gibt an, wie RDATA zu interpretieren ist
- TTL: wie lange der RR in einem Cache zwischengespeichert werden darf
- RDATA: die Daten des RR

Codierung von „Namen“

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
20	**7**								**"F"**							
22	**"R"**								**"A"**							
24	**"N"**								**"K"**							
26	**"E"**								**"N"**							
28	**2**								**"D"**							
30	**"E"**								**0**							
40	**3**								**"W"**							
42	**"W"**								**"W"**							
44	**1**	**1**	**20**													
46	**0**															

- Punkte sind die Trenner von Namen aber werden nicht codiert
- Anstelle von Punkten werden die Längen von Namenskomponenten codiert
- Die ersten zwei Bits einer Komponentenlänge dienen als Identifizierung von Pointern

„Namen“

- Domainkompenenten (Labels) sind auf maximal 63 Zeichen beschränkt
- Eine Domain (Name) darf aus maximal 255 Zeichen bestehen

www.franken.de

Label

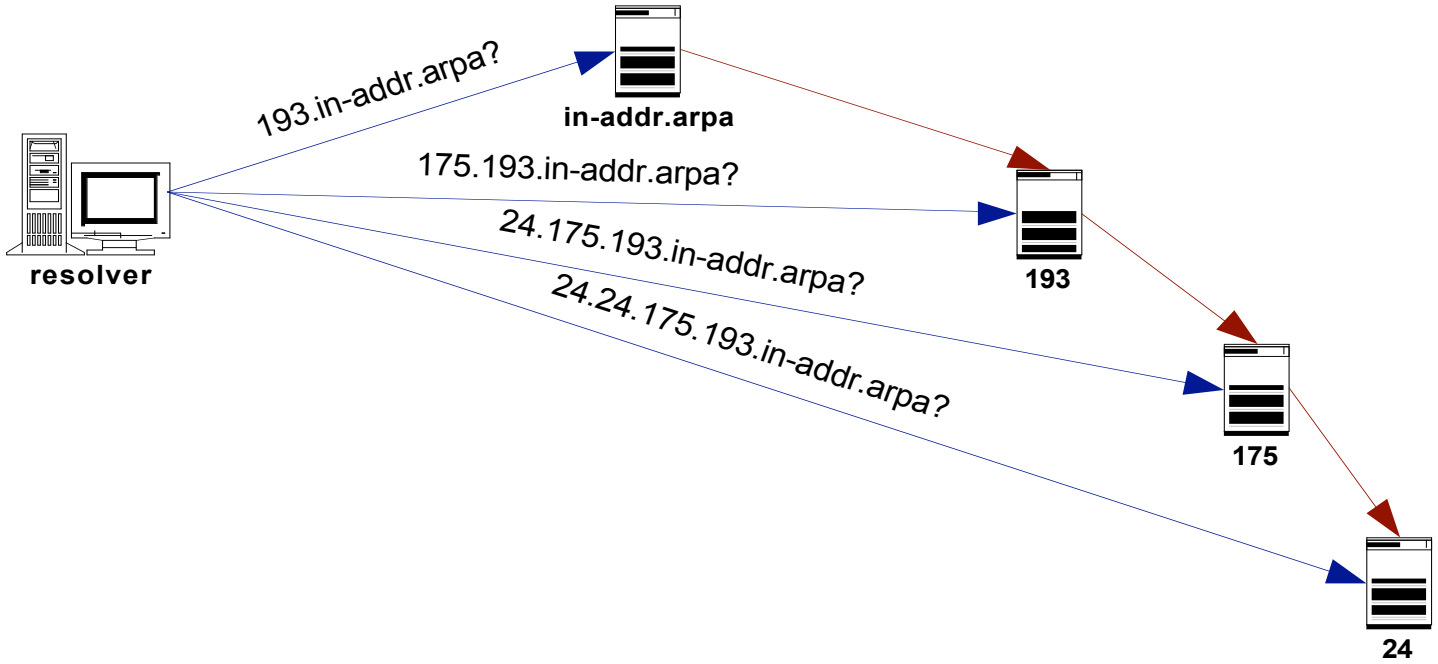
Domain

Typen von RRs

- A
  - 32bit IPv4 Adresse
- CNAME
  - anerkannter Name (Alias)
- NS
  - Nameserver
- SOA
  - Start einer Autorität
- MX
  - Mail Austauscher
- HINFO
  - Host Information
- TXT
  - Beliebiger Text
- AAAA
  - 128bit IPv6 Adresse
- PTR
  - Verweis auf anderen RR

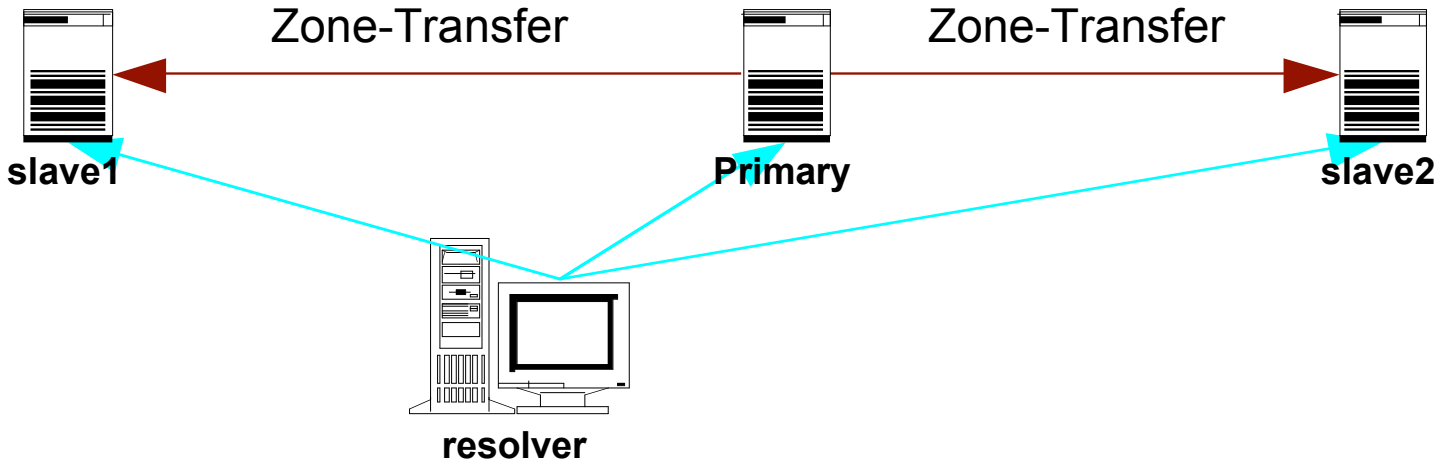
- KEY
- MB
- MD
- MF
- MG
- MINFO
- MR
- NAPTR
- NULL
- RT
- PX
- SIG
- SRV
- TKEY
- TSIG
- X25
- DNAME

Reverse Mapping



- Rückwärtsauflösung von 193.175.24.24
- IP zu Namen mittels der in-addr.arpa Zone
- 193.175.24.24 wird erfragt mittels 24.24.175.193.in-addr.arpa
- Ergebnis:
  IN PTR www2.franken.de.

Slave & Master Server



- Slave-Server „ziehen“ die komplette Zone vom Primary
- Clients können alle NS abfragen und erhalten von ihnen „authorative“ Antworten
- Ausfall eines Nameservers führt zu nahezu keiner Beeinträchtigung des DNS Dienstes
- Wenn Primary nicht nach aussen sichtbar ist, wird das Setup „Hidden Primary“ genannt

SOA Record

```
franken.de IN SOA  dns.franken.de.
   dnsmaster.franken.de. (
  2004112001 ; serial
        7200 ; refresh (2 hours)
        1800 ; retry (30 minutes)
     1209600 ; expire (2 weeks)
       20000 ; minimum (5 hours 33 mins 20 secs)
 )
```

- Wird von Slave-Servern benutzt um Zonen-Informationen zu erhalten, z.B. bzgl. Aktualisierungen
- Kann benutzt werden um Verantwortlichen einer Zone zu ermitteln
- @-Zeichen in SOA-Records wird durch Punkt codiert, local-part selbst darf daher keinen „.“ enthalten (dnsmaster@franken.de -> dnsmaster.franken.de)

MX Record

- Mailexchanger für Internet Mail
- Spezifiziert Hosts, die Mails für eine Domain entgegennehmen
- Mehrere MX Records pro Name möglich
- MTAs sollten die als MX gelisteten Rechner in der Reihenfolge ihrer Priorität kontaktieren

```
franken.de. IN   MX 10 laurel.franken.de.
franken.de. IN   MX 20 priscilla.franken.de.
```

- MX RRs dürfen nicht auf RRs zeigen, die auf andere RRs verweisen, so darf z.B. ein in einem MX referenzierter Hostname kein CNAME sein.

CNAME Record

- Verweis auf anderen (primären) Namen eines Hostnamens
- Ein Hostname darf ausser einem CNAME keine weiteren RRs haben
- MX oder NS Records dürfen nicht auf CNAMEs zeigen
- Ein CNAME darf nicht auf einen weiteren CNAME zeigen

```
www.franken.de   IN CNAME   www2.franken.de.
www2.franken.de. IN A       193.175.24.24
```

A Record

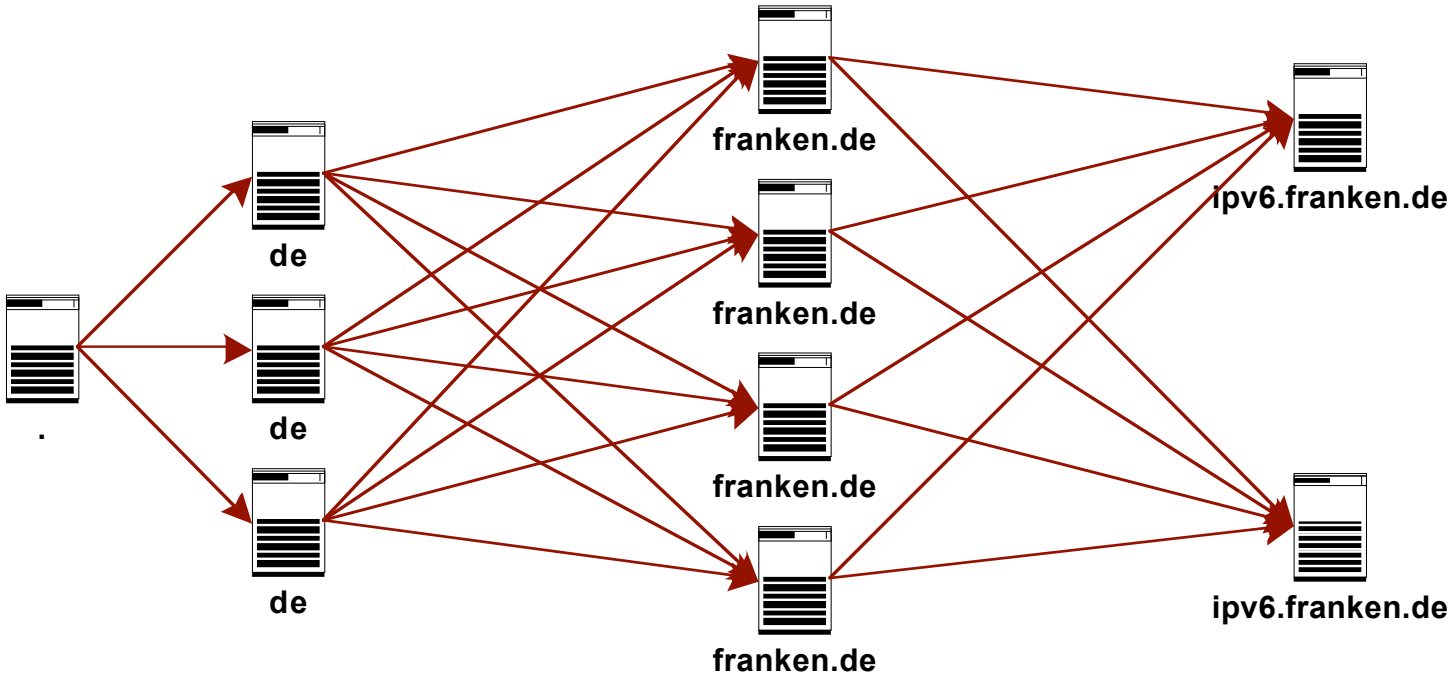
- Zuordnung einer IP-Adresse zu einem Host-/Domain-Namen
- Darf zusätzliche RRs enthalten, meist HINFO, TXT und MX
- Host-/Domain-Namen dürfen mehrere A RR besitzen

```
elvis.franken.de.   IN A     193.175.24.41
elvis.franken.de.   IN HINFO "PIII-650" "Linux"
elvis.franken.de.   IN MX    5 elvis.franken.de.
elvis.franken.de.   IN MX    10 rachael.franken.de.
elvis.franken.de.   IN MX    50 faui45.informatik.uni-
   erlangen.de.
```

NS Record

- Liefert verantwortliche Server für eine „Zone“
- DNS ist durch die Verknüpfung von Zonen aufgebaut, NS Records knüpfen diese zu einem gerichteten Graphen zusammen

```
ipv6.franken.de. IN NS ip6-r.franken.de.
ipv6.franken.de. IN NS dns.franken.de.
```



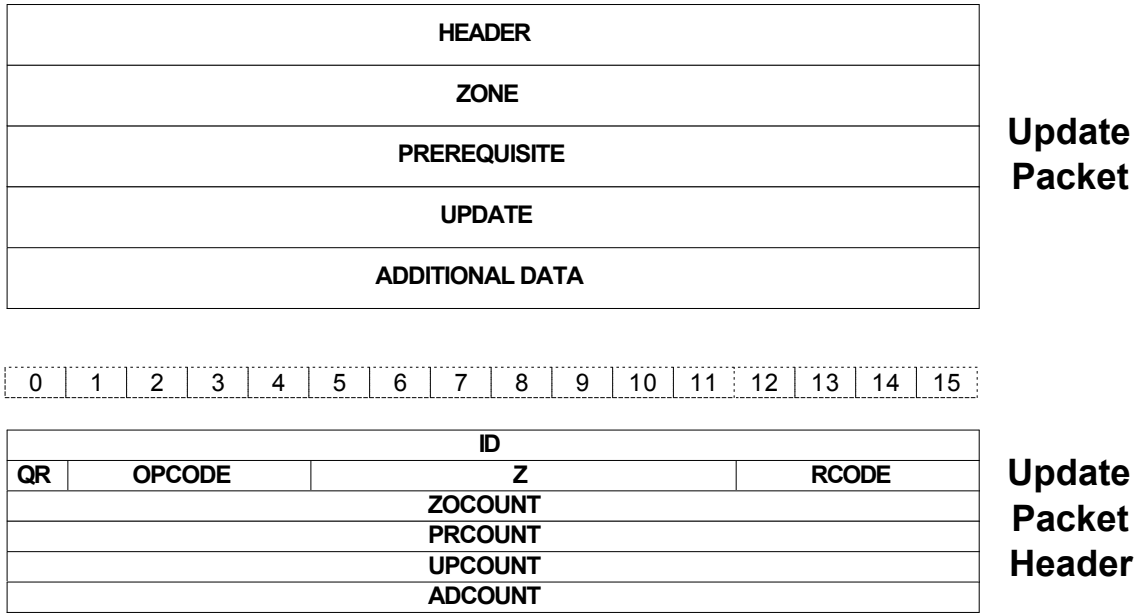
TTL

```
dns.franken.de. 43200 IN A 193.175.24.33
```

- Time To Live
  Zeitspanne (in Sekunden), die sich ein DNS-Cache den Resource-Record merken darf
- Jeder RR kann eigene TTL haben
- Hohe TTL verringert Traffic auf Nameservern
- Niedrige TTL ermöglicht häufige Änderungen ohne Inkonsistenzen
- Programme haben keine Möglichkeit über Standard-APIs („gethostbyname()“ etc) auf die TTL zuzugreifen und cachen Adressen daher häufig mit frei gewählten TTLs

Dynamic DNS

- Häufige oder automatisierte Änderungen von RRs in Zonen
- Änderungen werden nicht durch „reload“ der Konfiguration sondern durch eine Netzwerk-API vorgenommen (RFC 2136 „DNS Update“)
- Updates werden zu Transaction zusammengefasst



Dynamic DNS

- Update der Adressen von Hostnamen, deren Rechner über Dialin-Pools eingewählt sind
  - Vielzahl von Protokollen zwischen Client (Einwahlrechner) und DynDNS-Server
  - TTL von via DynDNS adressierten Rechnern sehr niedrig -> häufige DNS-Nachfragen nötig
  - „disconnect“ meist nicht feststellbar, daher keine Zuordnung von MX-RR sinnvoll
  - Trotzdem hervorragend geeignet für
    - Webserver auf Rechner an Dialup
    - SSH-Zugang auf Rechner an Dialup

Cache Poisoning

„Verschmutzung“ eines DNS Caches mit falschen Daten

- „Birthday Attack“
  - Fluten eines DNS-Caches mit Anfragen und eigenen Antworten wodurch nach einer gewissen Zeit die Antwort eine zur Anfrage passende ID hat
- Phase Space Analysis Spoofing
  - Versuch Regelmässigkeiten im Zufallszahlengenerator des Cache-Systems zu ermitteln um damit die ID von DNS-Anfragen ausrechnen zu können

Typische DNS-Probleme

- Inkonsistente Slaves
- Fehlerbehaftete Zonen
- Zu tief verlinkte RRs
- „Lame“ Server
- Administrative Fehler

Inkonsistente Slave-Server

Nicht alle Server liefern die gleichen Daten, einige Server liefern veraltete Informationen

- Update der Zone auf dem Master ohne die Serial-Number erhöht zu haben
- Slave-Server hat keine Berechtigung zum Zone-Transfer vom Master
  - Falsche ACL auf dem Nameserver
  - Probleme mit der Firewall
  - Slave hat neue IP-Adresse
- Zone-Konfiguration auf Masterserver enthält Fehler, wodurch der Server keine AA Antworten verschickt - Secondaries holen daraufhin die Zone nicht mehr vom Master

Fehlerbehaftete Zonen

- Allgemeine Syntax-Fehler
- Hostname mit CNAME und weiteren RRs
- Unzulässige Zeichen in Namen
- CNAME oder MX der wiederum auf CNAME oder MX zeigt
- „trailing dot“ bei FQDN vergessen wodurch „Echos“ entstehen (Beispiel: mail.franken.de.franken.de)
- Zone enthält essentielle Records mit nicht erreichbaren IP-Adressen (z.B. „IN NS 192.0.2.1“)

Zu tief verlinkte Zonen

- Durch die Baumstruktur von DNS können zu tief verschachtelte Strukturen gebildet werden, die zu Timeout-Problemen führen

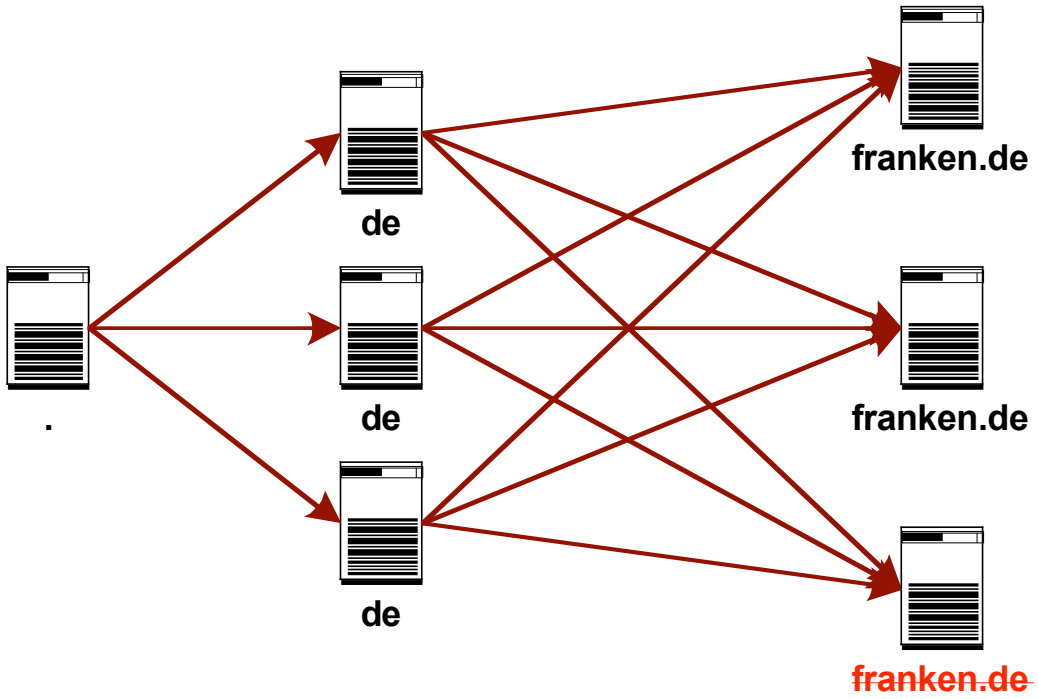
```
www.domain1.de IN CNAME www2.foo.dom2.de
foo.dom2.de    IN NS dns.bar.domain3.de
bar.domain3.de IN NS dns.domain4.de
domain4.de     IN NS dns.domain4.de
```

- Glue-Records verwenden!

```
www.domain1.de IN CNAME www2.foo.dom2.de
www2.foo.dom2.de IN A 192.0.2.1
```

Lame Server

- Delegation zu Nameservern, die Zonen nicht (mehr) führen



Administrative Fehler

- RAM-Knappheit
  DNS-Server dürfen nicht in Swapspace ausgelagert werden
- Alle DNS-Server einer Zone in gleichem Subnetz
- Zu hohe TTL vor Änderung von wichtigen Daten
- Serial-Number der Zone zeitweise „kaputt“

Tools: dig

```
#dig www.franken.de

; <<>> DiG 9.2.1 <<>> www.franken.de
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 4804
;; flags: qr aa rd ra; QUERY: 1, ANSWER: 2, AUTHORITY: 4, ADDITIONAL: 5

;; QUESTION SECTION:
;www.franken.de.                        IN      A

;; ANSWER SECTION:
www.franken.de.         43200   IN      CNAME   www2.franken.de.
www2.franken.de.        43200   IN      A       193.175.24.24

;; AUTHORITY SECTION:
franken.de.             43200   IN      NS      ns.rrze.uni-erlangen.de.
franken.de.             43200   IN      NS      dns.franken.de.
franken.de.             43200   IN      NS      arbi.informatik.uni-
    oldenburg.de.
franken.de.             43200   IN      NS      secondary.rrze.uni-
    erlangen.de.

;; ADDITIONAL SECTION:
ns.rrze.uni-erlangen.de. 52273  IN      A       131.188.3.2
dns.franken.de.         43200   IN      A       193.175.24.33
```

Tools: dnstracer

```
#dnstracer -c -o -s . franken.de
Tracing to franken.de via A.ROOT-SERVERS.NET, timeout 15 seconds
A.ROOT-SERVERS.NET [.] (198.41.0.4)
 |\___ B.DE.NET [de] (194.97.99.44)
 |      |\___ ns.rrze.uni-erlangen.de [franken.de] (131.188.3.2)
 |      |\___ secondary.rrze.uni-erlangen.de [franken.de] (131.188.3.4)
 |      |\___ arbi.informatik.uni-oldenburg.de [franken.de] (134.106.1.7)
 |       \___ dns.franken.de [franken.de] (193.175.24.33)
 |\___ C.DE.NET [de] (193.159.170.187)
 |      |\___ secondary.rrze.uni-erlangen.de [franken.de] (131.188.3.4)
 |      |\___ ns.rrze.uni-erlangen.de [franken.de] (131.188.3.2)
 |      |\___ arbi.informatik.uni-oldenburg.de [franken.de] (134.106.1.7)
 |       \___ dns.franken.de [franken.de] (193.175.24.33)
 |\___ I.DE.NET [de] (206.65.170.100)
 |      |\___ dns.franken.de [franken.de] (193.175.24.33)
 |      |\___ ns.rrze.uni-erlangen.de [franken.de] (131.188.3.2)
 |      |\___ secondary.rrze.uni-erlangen.de [franken.de] (131.188.3.4)
 |       \___ arbi.informatik.uni-oldenburg.de [franken.de] (134.106.1.7)
```

Neue Anwendungen

- RBL
  Realtime Blocking Lists
- SSH-Hostkey Verifikation über fingerprint resource records in DNS
- SPF „Sender Permitted From“ erlaubt es Domain-Inhabern anzugeben, welche Rechner die Domain im „From“ verwenden dürfen

DNS-Roots

A	VeriSign Global Registry	Dulles VA
B	Information Sciences Institute	Marina Del Rey CA
C	Cogent Communications	Herndon VA; Los Angeles; New York City; Chicago
D	University of Maryland	College Park MD
E	NASA Ames Research Center	Mountain View CA
F	Internet Systems Consortium, Inc. (ISC)	Ottawa; Palo Alto; San Jose CA; New York City; San Francisco; Madrid; Hong Kong; Los Angeles; Rome; Auckland; Sao Paulo; Beijing; Seoul; Moscow; Taipei; Dubai; Paris; Singapore; Brisbane; Toronto; Monterrey; Lisbon; Johannesburg; Tel Aviv; Jakarta; Munich; Osaka; Prague
G	U.S. DOD NIC	Vienna VA
H	U.S. Army Research Lab	Aberdeen MD
I	Autonomica/NORDUnet	Stockholm; Helsinki; Milan; London; Geneva; Amsterdam; Oslo; Bangkok; Hong Kong; Brussels; Frankfurt; Ankara; Bucharest; Chicago; Washington DC; Tokyo; Kuala Lumpur
J	VeriSign Global Registry Services	Dulles VA (2 locations); Mountain View CA; Seattle WA; Amsterdam; Atlanta GA; Los Angeles CA; Miami; Stockholm; London; Tokyo; Seoul; Singapore; Sterling VA (2 locations, standby)
K	Reseaux IP Europeens - Network Coordination Center	London (UK); Amsterdam (NL); Frankfurt (DE); Athens (GR); Doha (QA); Milan (IT); Reykjavik (IS); Helsinki (FI); Geneva (CH)
L	IANA	Los Angeles
M	WIDE Project	Tokyo; Seoul (KR); Paris (FR)

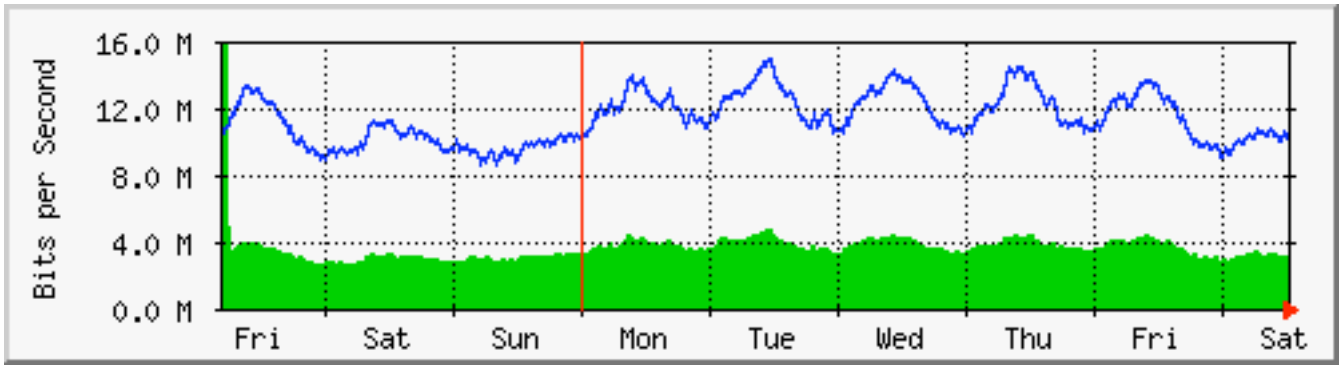
Statistik

- Im Juli 2004 ungefähr 285 Millionen Hosts
- Top20 Hostnamen:

1	www	11	pc1
2	mail	12	pc2
3	cpe	13	server
4	ns	14	broadcast
5	ns2	15	pc3
6	ns1	16	pc4
7	ftp	17	pc5
8	gw	18	gateway
9	router	19	www2
10	smtp	20	pc6

Statistik

- Traffic auf h.root-servers.net:



- 247 ccTLDs (Country Code Toplevel Domains, siehe ISO-3166-1)
- 14 Generic TLDs
- 1 Infrastructure Domain (.arpa)

Generische TLDs

The .aero domain is reserved for members of the air-transport industry

The .biz domain is restricted to businesses

The .com domain

The .coop domain is reserved for cooperative associations

The .info domain

The .museum domain is reserved for museums

The .name domain is reserved for individuals

The .net domain

The .org domain is intended to serve the noncommercial community

The .pro domain is being established; it will be restricted to credentialed professionals and related entities

The .gov domain is reserved exclusively for the United States Government.

The .edu domain is reserved for postsecondary institutions accredited by an agency on the U.S. Department of Education's list of Nationally Recognized Accrediting Agencies

The .mil domain is reserved exclusively for the United States Military.

The .int domain is used only for registering organizations established by international treaties between governments.

Links

- Internet Systems Consortium http://www.isc.org
  Quelle für BIND Nameserver
- dnsreport http://www.dnsreport.com/
  Online DNS-Problem Tester
- dnstracer
  http://www.mavetju.org/unix/dnstracer.php
  Kommandozeilen-Tool um DNS Baum zu Verfolgen
- RFC 1034 „Domain names - concepts and facilities “
  ftp://ftp.rfc-editor.org/in-notes/rfc1034.txt
- RFC 1035 „Domain names - implementation and specification“
  ftp://ftp.rfc-editor.org/in-notes/rfc1035.txt
- RFC 1912 „Common DNS Operational and Configuration Errors“
  http://www.rfc-editor.org/rfc/rfc1912.txt