AS112

Elmar K. Bins <bins@denic.de>

DENIC Networking

29. September 2005

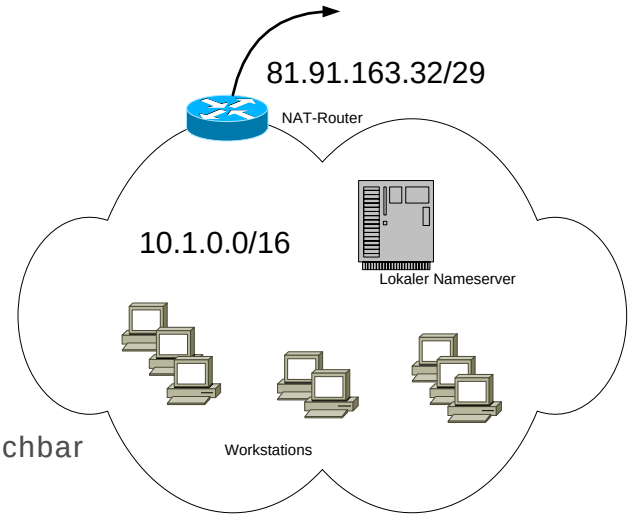
Überblick

- Die Welt seit RFC 1918
- „Leakende“ Queries und Updates
- Lastverringerung durch Delegation
- „Leakende Queries“ und Updates jetzt
- Lastverteilung per Anycasting – Projekt „AS112“
- BGP Anycasting
- DENIC – Query- und Updateraten
- Was kann ich im eigenen Netz tun?
- Optionen für (Access-) ISPs
- Oops!
- Quellen

Die Welt seit RFC 1918 (1996)

Intern RFC-1918 („private“) Adressen

- 10/8, 172.16/12 und 192.168/16
- Ausreichend Adreßraum
- Flexibilität, keine Bürokratie



Umsetzung nach außen via NAT

- Erreichbarkeit des „Internet“
- interne Knoten von außen nicht erreichbar

Irgendwas geht immer schief

- Meistens ist es (Reverse-) DNS

Gern gemachte Fehler

Pakete mit Source-Adressen aus RFC-1918 tauchen im Internet auf

- Sauberes NAT am Router
- Filtern ausgehender IP-Pakete (egress filtering)

DNS-Abfragen im internen Netz

- Interner DNS-Server löst interne Domainnamen auf
- Leider wird Reverse oft vergessen
- Workstations senden oft selbst DNS-Queries in die Welt

Ein DNS-Query „leakt“ ins Internet

Typischer DNS-Reverse-Query für RFC 1918-Adreßraum

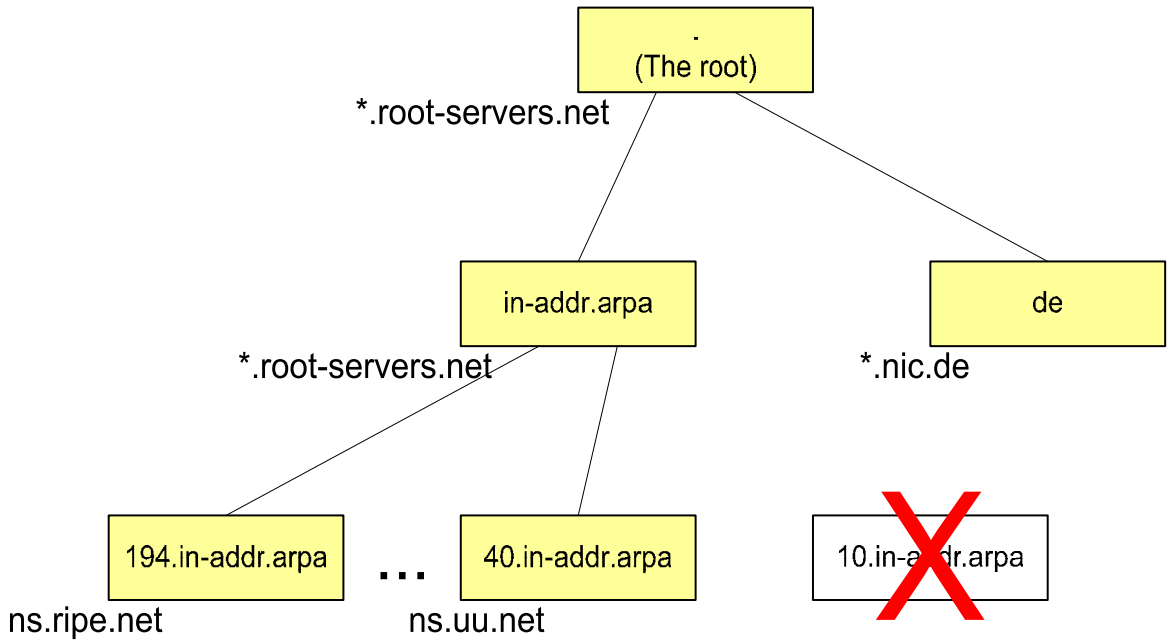
```
193.192.100.120.33570 > a.b.c.d.53: 41091 [1au] PTR? 6.1.1.10.in-addr.arpa. (53)
```

- Source-Adresse ist ok: **193.192.100.120** (dank NAT)
- Aber die DNS-Anfrage! **PTR? 6.1.1.10.in-addr.arpa.**

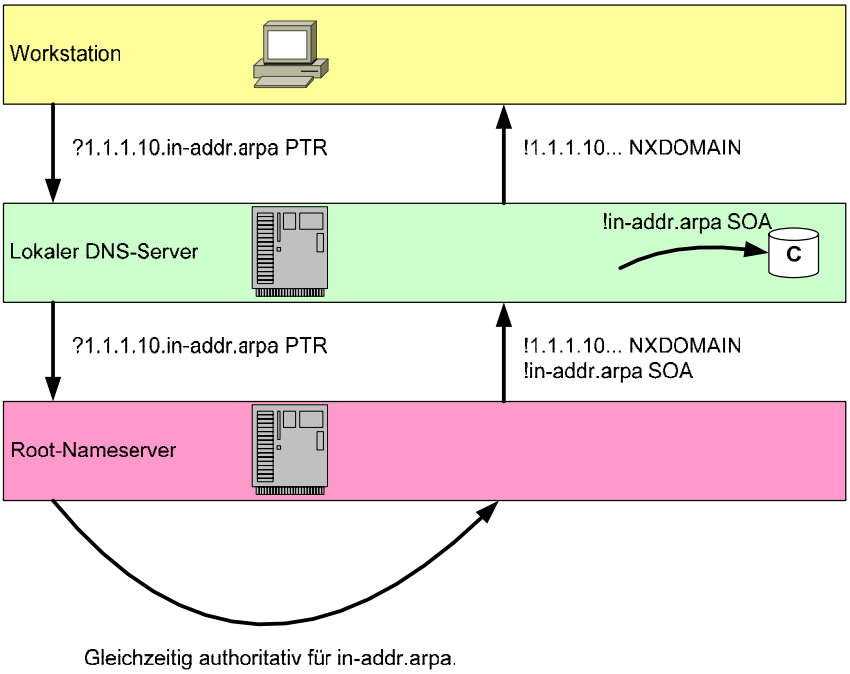
Darauf kann es aber keine Antwort geben

- Anfrage nach interner Infrastruktur
- Diese Infrastruktur ist „draußen“ natürlich unbekannt
- Sie interessiert genau genommen auch niemanden

Der DNS-Baum (vor 2000)



Ablauf eines Query



Dynamische Updates

Als wäre das noch nicht genug

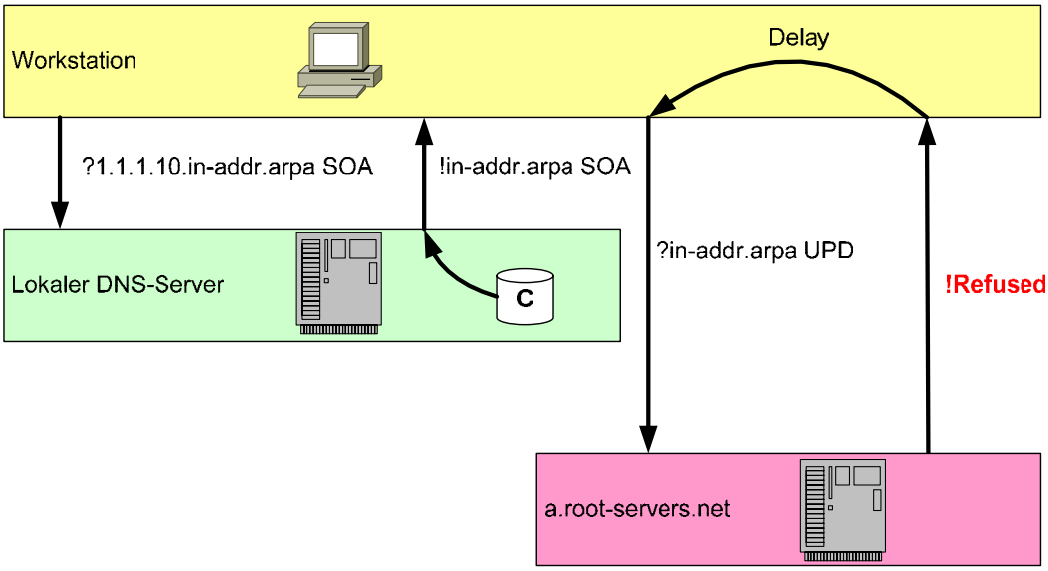
```
81.13.158.250.27363 > a.b.c.d.53: 3035 update [1n] SOA? 10.in-addr.arpa.
```

- Dynamische Updates sind toll – in der eigenen Infrastruktur
- Müssen ignoriert werden, das kostet CPU (je nach Software)

Wo kommen die Updates her?

- DHCP vergibt Adressen und trägt sie ins DNS ein (A, PTR)
- Einige Workstations tun das auch gern mal selber
- Verursacher: Meist Windows (Voreinstellung „update DNS“)
- Um Mitternacht geht‘s **richtig** rund (DHCP lease renew)

Ablauf eines Update (-versuchs)



- a.root-servers.net (aus dem SOA) kriegt die ganze Update-Last ab

Lastverringerung durch Delegation

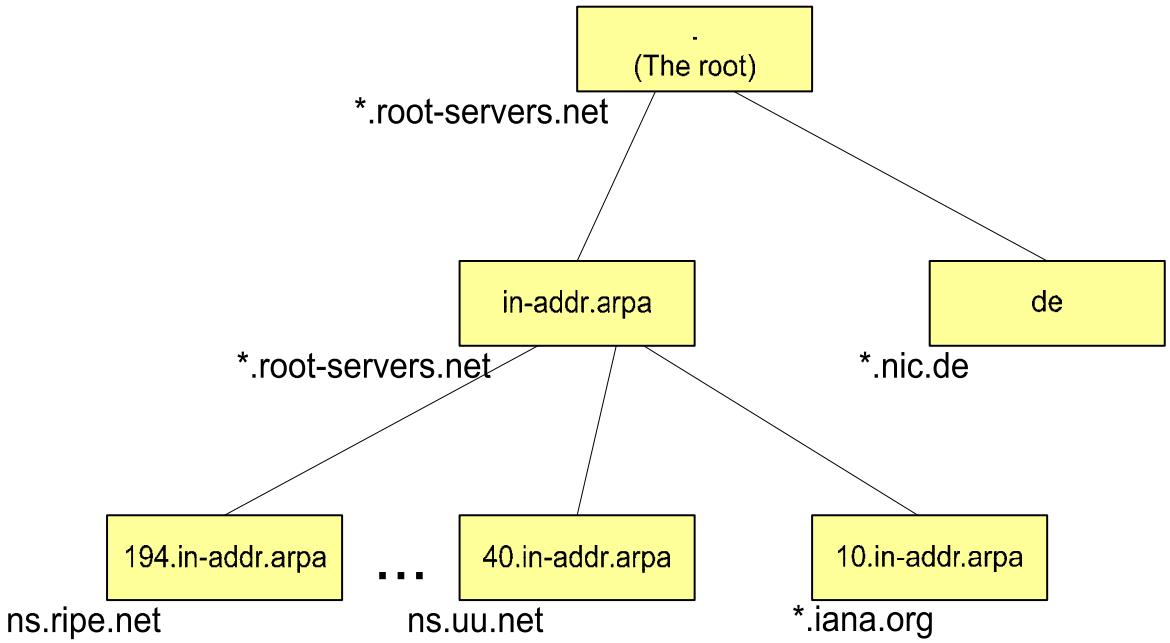
Hohe überflüssige Last auf den Root-Servern

- Jan. 2001 etwa 12000 Queries/s auf a.root (5000 auf f.root)
- Nov. 2002 etwa 1200 Updates/s auf a.root
- Tendenz steigend

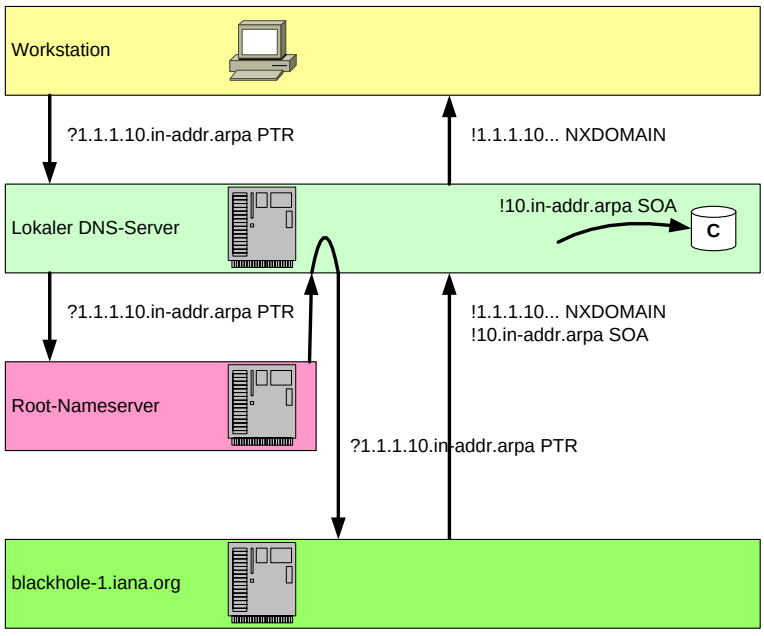
Entlastung durch Delegation

- Einzelne RFC-1918-Reverse-Zonen werden erzeugt und delegiert
- Delegationen an blackhole-1/blackhole-2.iana.org
- prisoner.iana.org im SOA, der dann die Updates bekommt
- Keine Updates mehr an a.root-servers.net

Der DNS-Baum nach der Delegation

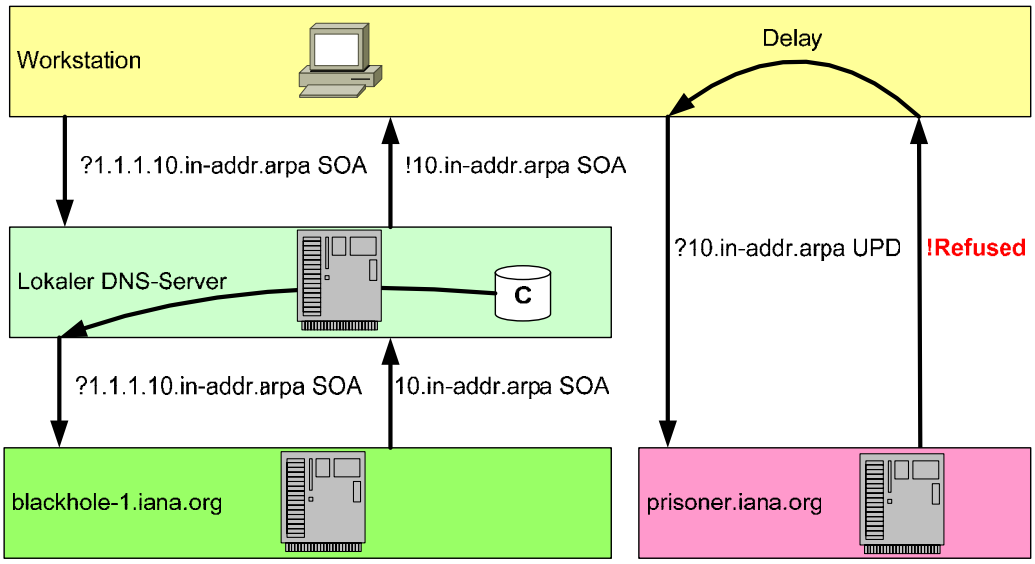


Ablauf eines Query jetzt



- Weitere Reverse-Queries für *.10.in-addr.arpa beginnen dank Cache gleich bei blackhole-1.iana.org

Ablauf eines Update jetzt



- Die Updates gehen nicht mehr an einen Root-Nameserver
- Ziel erreicht!

Lastverteilung durch Anycasting

Hohe Last jetzt auf den IANA-Servern

- Ziel erreicht?
- Die Workstations versuchen es weiterhin nach einem Delay erneut
- Belastung der IANA-Server steigt stetig, denn das Internet wächst

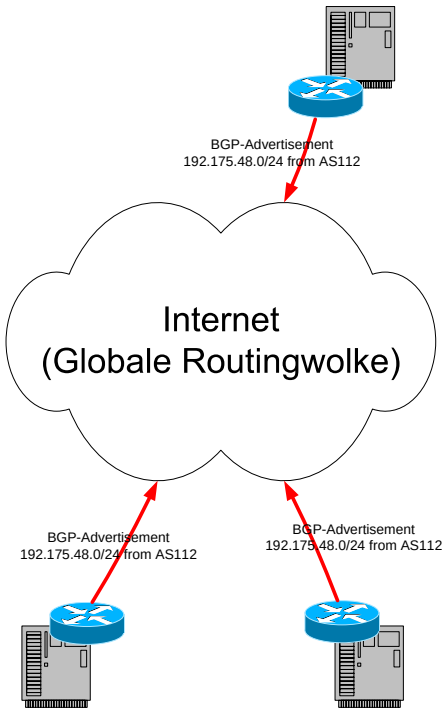
Projekt „AS112“ wird geboren

- Die Community findet eine Lösung für die Lastverteilung
- Man vergibt AS112 und 192.175.48.0/24 für das Projekt
- Mehrere Teilnehmer hosten geeignete Systeme rund um die Welt
- Die Lastverteilung geschieht mittels BGP-Anycasting

BGP-Anycasting

BGP-Anycasting

- An mehreren Stellen der Welt wird das gleiche Netz (vom gleichen AS aus) advertised
- Im globalen BGP existieren mehrere Pfade zum Ziel
- Der eigene Router wählt einen „besten“ Pfad, je nach eigener „Weltsicht“
- Fällt ein Standort aus, verschwindet der zugehörige Pfad, ein anderer Standort ist nun der „beste“ für uns



Verfügbare Pfade aus Sicht der DENIC (AS8763)

Mehrere Pfade für 192.175.48.0/24 (es gibt noch viel mehr)

```
...
Paths:
...
8220 16150 112          AS112-Instanz bei „Port 80 AB“
12312 3257 112          AS112-Instanz bei „Tiscali intl.“
112 (used)              AS112-Instanz bei der DENIC
...
```

Nun fällt unsere eigene Instanz aus (*rrrums*)

```
...
Paths:
...
8220 16150 112          AS112-Instanz bei „Port 80 AB“
12312 3257 112 (used)   AS112-Instanz bei „Tiscali intl.“
...
```

- Wir haben einen neuen „besten“ Pfad
- Das Projekt skaliert und ist recht ausfallsicher
- Schlecht vorhersagbar, wieviel Traffic eine Instanz abbekommt

Kontrollmöglichkeiten für‘s BGP

Einschränken der Sichtbarkeit durch BGP-Communities

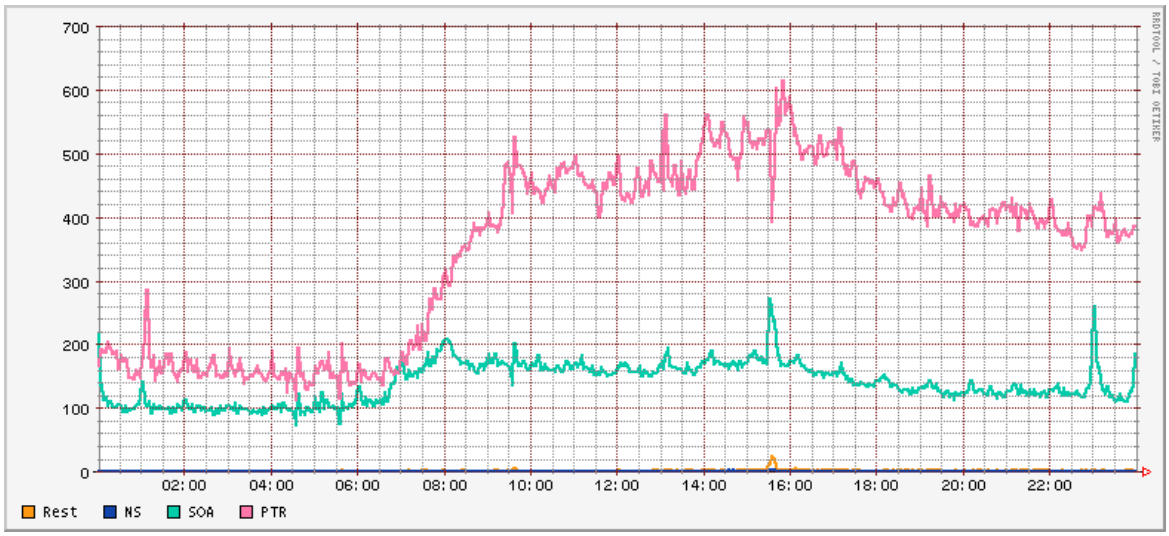
- Prefix 192.175.48.0/24 „no-export“ an BGP-Partner geben
- Damit ziehen wir nur den Traffic der Partner selbst an
- Damit erwischen wir **nicht** ihre BGP-Kunden!

Einschränken der Sichtbarkeit durch Selektion

- Prefix 192.175.48.0/24 nur an Peers, nicht an Transit-ISPs geben
- Damit ziehen wir fast nur den Traffic der Peers und ihrer Kunden an
- Wenig Traffic darüber hinaus (andere Pfade sind besser)
- Damit kriegen wir immer noch eine Menge Anfragen

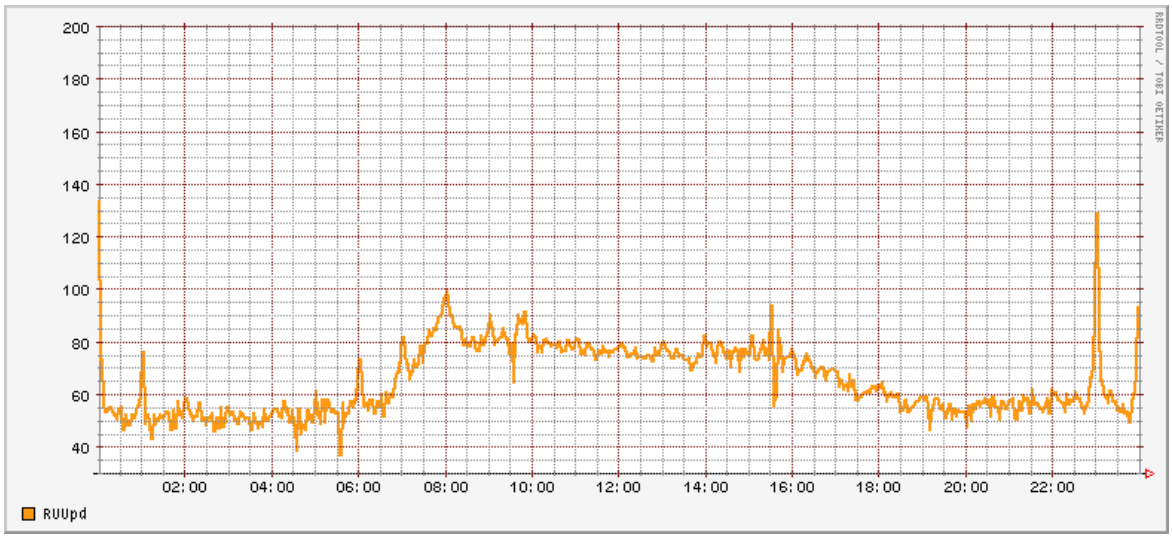
Anzahl Queries/s für Montag

- Wochenanfang – die Kurve bleibt zum Dienstag hin oben



Anzahl Updates/s (wieder vom Montag)

- man erkennt am Rand noch die „Mitternachtsupdates“



Soll ich da mitmachen?

Warum nimmt die DENIC am Projekt „AS112“ teil?

- Unser „Job“ ist DNS
- Wir fühlen uns als Teil der Community
- „Tue Gutes und rede darüber“
- (außerdem sind wir neugierig)

Entlastung des Internet von überflüssigen Anfragen

- Geht auf verschiedenen Wegen
- ISPs haben andere Anforderungen und Möglichkeiten

Was kann ich im eigenen Netz tun?

Lokale Nameserver „hilfreich“ konfigurieren

- Lokale Nameserver mit den RFC-1918-Reversezonen bestücken
  - `(10.in-addr.arpa, 168.192.in-addr.arpa, 16-31.172.in-addr.arpa)`
- Das sorgt für Abhilfe bei den Reverse-Queries
- Außerdem erhält man funktionierendes Reverse-Mapping ;-)

Klassische Network Security

- An Firewall/Router DNS-Zugriff „nach draußen“ verbieten
- Das verhindert zufällige Zugriffe und Updates an externe Server

Optionen für (Access-) ISPs

Kunden sind oft nicht kontrollierbar, Filter nicht durchsetzbar

- „Wilde“ DNS-Requests können nicht gefiltert werden
- Die den Kunden angebotenen DNS-Server können aber mit den RFC-1918-Zonen bestückt werden. Das hilft schon ein bißchen...

Eine lokale AS112-Instanz

- Die meisten ISPs verwenden ohnehin BGP im Backbone
- Eine AS112-Instanz (DNS-Server und BGP-Server, z.B. Quagga) ist schnell aufgesetzt und recht betriebsstabil
- 192.175.48.0/24 muß zu anderen ISPs (ausgehend) gefiltert werden
- Voilà - Kundentraffic zu 192.175.48.0/24 landet auf dem eigenen AS112-Server (Requests **und** Updates)

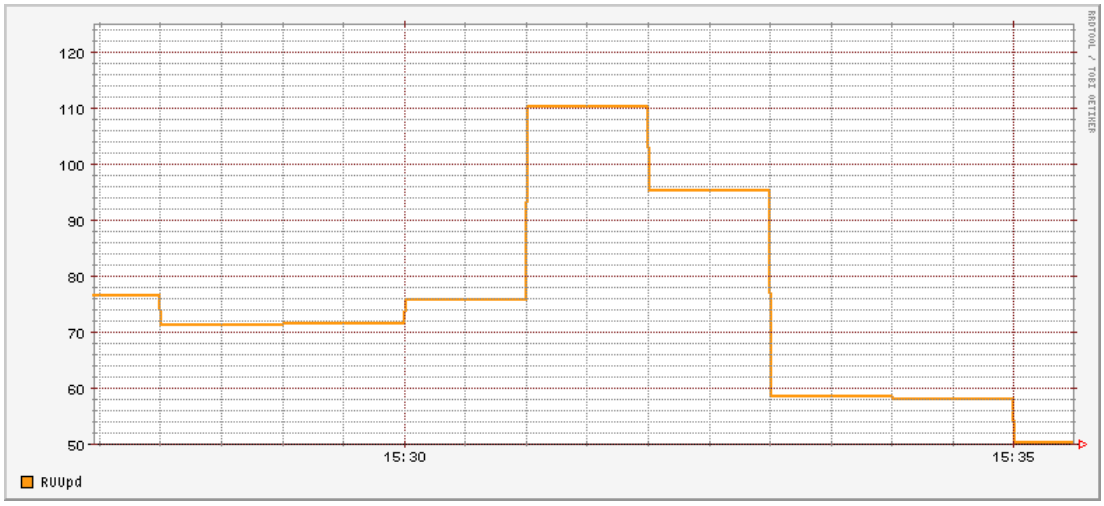
Engagement im AS112-Projekt (für Ambitionierte)

Die eigene („globale“) AS112-Instanz

- Mit dem AS112-Team koordinieren (u.a. RaDB-Pflege)
- Die ausgehenden BGP-Filter entfernen
- Sicherstellen, daß der AS112-Server die Last verträgt
- **Oops! Zuviel!**

Da ist uns was passiert...

- Netz wurde kurzfristig an Transit-ISP weitergegeben (15:31-15:32)
- Limitiert durch den Server (100% Last)
- Neue Maschine ist bestellt ;-)



Fragen?

(Grafik schamlos geklaut bei: RIPE NCC)

Quellen

AS112 project / DNS for RFC-1918 reverse

AS112-Projekthomepage
http://www.as112.net/

CAIDA-Präsentationen 2001 und 2003
http://www.caida.org/outreach/presentations/ietf0112/dns.damage.html
http://www.caida.org/outreach/presentations/2003/wiapp03/sdu.wiapp03.slides.pdf

Anycast

Deploying IP Anycast (CMU, Kevin Miller)
http://www.net.cmu.edu/pres/anycast/

DNS-Resolvers

Configuration Issues... (Proposal, Mark Andrews, ISC)
http://www.ietf.org/internet-drafts/draft-andrews-full-service-resolvers-00.txt