

IPv6 in der Praxis

Felix von Leitner

CCC Berlin

`felix@ccc.de`

Dezember 2000

Agenda

1. Autoconfiguration, Neighbor Discovery
2. Router Discovery, Router Renumbering
3. BSD socket API Erweiterungen
4. Multicast Listener Discovery
5. DNS, FTP
6. Spezialitäten

Medien

IPv6 ist definiert für Transport über

1. Ethernet (RFC 2464)
2. Token Ring (RFC 2470)
3. Frame Relay (RFC 2590)
4. FDDI (RFC 2467)
5. ATM (RFC 2492)
6. ARCnet (RFC 2497)

EUI-64

[*] IPv6: enable EUI-64 token format

[*] IPv6: disable provider based addresses

EUI-64 ist eine IEEE-Norm für einen globalen 64-bit Global Identifier. IPv6 hat vorher 48-bit Adressen vorgesehen. Problem: Prefixlänge ändert sich von 80 auf 64 Bits, der Standard mußte angefaßt werden.

IEEE definiert das Mapping von 48-bit Ethernet Adresse zu EUI-64:

34-56-78-9A-BC-DE -> 36-56-78-FF-FE-9A-BC-DE

(Bit 7 ist das „Universal/Local“ Bit und im Ethernet gewöhnlich 0).

A Compact Representation of IPv6 Addresses (RFC 1924)

Base85-Encoding mit

'0'..'9', 'A'..'Z', 'a'..'z', '!', '#', '\$', '%', '&',
'(', ')', '*', '+', '-', ';', '<', '=', '>', '?', '@',
'^', '_', '`', '{', '|', '}' und '~'.

Beispiel:

1080:0:0:0:8:800:200C:417A ->

21932261930451111902915077091070067066 ->

4)+k&C#VzJ4br>0wv%Yp

Das RFC ist vom 1. April 1996...

IPv6 Bootstrap

1. `::1/128` ist die Host Local Adresse (analog 127.0.0.1)
2. Die Link-Local Adresse ist `fe80:: [EU-64]`.
D.h. für `00:11:22:33:44:55` ist sie `fe80::0211:22FF:FE33:4455`.
3. Per Neighbor Solicitation (wie ARP) fragt der Host nach, ob diese Adresse bereits vergeben ist (würde ein Neighbor Advertisement nach sich ziehen).
4. Mit der Link-Local Adresse kann der Host kommunizieren, d.h. sich z.B. das Kernel-Image holen oder mit Routern sprechen.
5. Der Host sendet Router Solicitations.

IPv6 Bootstrap

6. Die Router senden Prefixe mit Lease Timeouts. Wenn kein Router antwortet, muß der Host DHCPv6 probieren.

7. Die Router melden auch MTU und Hop-Count.

Pro Link gibt es mehrere Prefixe! Der richtige wird bei der Path MTU Discovery gefunden.

Router Advertisements beinhalten link-layer Adressen, d.h. kein zusätzliches ARP o.ä. ist nötig.

Die gesamte Prozedur benötigt keinerlei Konfiguration!

IPv6 Neighbor Discovery (RFC 2461)

- Ähnlich ARP, aber benutzt Multicast statt Broadcast!
- Die Multicast-Gruppe berechnet sich aus der IP-Nummer
- Erkennung von doppelt vergebenen IPs
- Neighbor Unreachability Detection hilft gegen stale ARP-Caches.
- Hop-Count = 255, d.h. keine Pakete von außen.
- Neighbor Discovery läuft auf ICMP-Ebene (ARP ist nicht so generisch).

Resource Discovery per Multicast

Bei IPv6 Multicast-Adressen sind 4 Bits für Scope reserviert.

So kann Resource Discovery feingetuned werden: Für Router benutzt man als Scope „Link“ , für DNS oder Web-Proxy würde man „Site“ benutzen.

Multicast-IP sind reserviert für all-hosts, all-routers und Router nach gewünschtem Routing-Protokoll.

Nachteil von Multicast: UDP; die Protokolle basteln sich ihre eigene Fragmentierung.

Router Selection

IPv6 Hosts haben einen Cache von Adressen und dazugehörigen next-hop Routern.

Router sagen in den Advertisements, ob sie default Router sind.

Hosts benutzen die Router bei unbekanntem Adressen round-robin. Nicht zuständige Router schicken einen Redirect.

Router Renumbering (RFC 2894)

Router können sich gegenseitig administrative Nachrichten schicken. U.a. können Prefixe gelöscht oder hinzugefügt werden.

Außerdem kann der Border Router ein Renumbering auslösen, das von den anderen Routern kaskadiert in alle Subnets weitergereicht wird.

Router Alert (RFC 2711)

Manche Pakete (z.B. RSVP oder Active Networks) haben zwar ein spezielles Ziel, aber sind auch für Router auf dem Weg wichtig.

Bei IPv4 müssen Router daher höhere Protokollebenen parsen.

Bei IPv6 gibt es die Router Alert Paketoption, mit der man Pakete markieren kann.

IPv6-Nummern in URLs (RFC 2732)

Da IPv6-Nummern Doppelpunkte benutzen, kollidieren sie mit URLs, die den Port durch einen Doppelpunkt abtrennen.

RFC 2732 spezifiziert daher folgende Syntax:

```
http://[1080::8:800:200C:417A]:80/index.html
```

Reservierte Anycast-Adressen

Anycast-Adressen sehen sind Adressen mit einem EUI-64 Wert von `fdff:ffff:ffff:ff80` bis `fdff:ffff:ffff:ffff`.

Davon ist momentan lediglich die ID 7E vergeben (für Mobile IP Agents).

Ich verstehe nicht, wieso noch niemand auf die Idee gekommen ist, hier Einträge für DNS, Email-Relay und Proxy-Server zu vergeben.

IPv6 socket API

Aus struct sockaddr_in wird struct sockaddr_in6.

```
struct sockaddr_in6 {
    uint8_t          sin6_len;          /* length of this struct */
    sa_family_t      sin6_family;      /* AF_INET6 */
    in_port_t        sin6_port;        /* transport layer port */
    uint32_t         sin6_flowinfo;    /* IPv6 flow information */
    struct in6_addr  sin6_addr;        /* IPv6 address */
    uint32_t         sin6_scope_id;
};
```

Aus gethostbyname wird gethostinfo, aus gethostbyaddr wird getaddrinfo. Und es gibt neue Adress-Test-Makros.

sin6_scope_id

Link Local IPv6 Nummern brauchen hier die Nummer des Netz-Interfaces (abfragbar mit `if_nametoindex` und `if_indextoname`), sonst schlagen `sendto`, `connect` und `bind` fehl.

RFC 2553 spezifiziert, daß eventuell auch für Site Local Adressen ein Interface eingetragen werden muß, überläßt die Details aber der Implementation.

Dieses Problem beißt alle möglichen schnell portierten Applikationen, wo dem User keine Möglichkeit gegeben wird, das Interface anzugeben (z.B. `ssh`), oder wo Daten per UDP oder separater TCP-Verbindung zurück übertragen werden (DNS, FTP).

Multicast Listener Discovery

Anstelle von IGMP benutzt IPv6 einen Subtyp von ICMPv6.

Die Spezifikation ist auf dem Niveau von IGMPv2, nicht IGMPv3 (keine ACLs).

Die jüngere Multicast-Routing Forschung ist daher nur teilweise übertragbar.

DNS und IPv6

1. Ansatz: „AAAA“ statt „A“ Records (RFC 1886). Reverse Lookup von ::1 über 1.0.0.0...0.0.ip6.int.

Problem: Keine Bit-Genauigkeit, keine Unterstützung für Renumbering.

2. Ansatz: Bitlabels (RFC 2673), DNAME (RFC 2672), „A6“ (RFC 2874). Reverse Lookup über ip6.arpa. Alle von einem Verrückten vom Fermilab und einem Microsoftie.

Problem: Unverständlich, widerlich, abstoßend, Brechreiz induzierend.

BIND8 kann IPv6 mit widerlichen Patches. BIND9 kann IPv6 mit DNAME, ist aber noch widerlicher. djbdns kann mit meinem Patch AAAA Records aber kein DNAME. Zusammenfassung: DNS muß sterben.

FTP und IPv6

Früher (RFC 959):

PORT 127,0,0,1,4,4

PASV -> 227 text (127,0,0,1,128,61)

Jetzt (RFC 2428):

EPRT |2|::1|1031|

EPSV -> 229 text (|||1234|)

Jumbograms (RFC 2675)

Jumbograms sind Pakete mit mehr als 65535 Bytes Daten. Jumbogramme sind optional und gedacht für Interfaces mit sehr hoher Bandbreite.

Das Ziel war, IPv6 innerhalb von Supercomputern zu benutzen.