IP ADVANCED

Problemstellung:

- Wie kann das Problem knapper IP-Adressen gelöst werden?
- Wie kann Multicast im Internet realisiert werden?

Lernziel:

Die Teilnehmer sollen die vorgestellten Systeme (NAT, IPv6, Multicast-Verfahren) erklären können.

Inhalt:

- Network Address Translation (NAT)
- IP Version 6
- IP Multicast

Network Address Translation (NAT)

Diese als "unsauber" (bzw. als "Hack") bezeichnete Methode erlaubt es, Rechner innerhalb von Intranets (abgeschlossene Netzwerke z.B. von Firmen) mit IP-Adressen zu betreiben, die nicht global eindeutig sind. Für die Kommunikation innerhalb eines Intranets sind lokal eindeutige Adressen ausreichend. Allerdings soll meist auch mit Rechnern im Internet kommuniziert werden, so dass eine spezielle Lösung notwendig wird.

NAT ermöglicht es (meist Firmen), relativ große Intranets mit relativ wenigen, global eindeutigen IP-Adressen zu betreiben. Innerhalb des Intranets wird meist die zur Klasse A gehörige Netzadresse 10 verwendet, die ursprünglich dem Arpanet zugewiesen war und heute nicht mehr benutzt wird. Bei NAT wird ein Intranet über einen speziellen Router, eine sogenannte *NAT-Box* mit dem Internet verbunden. Der NAT-Box wird eine kleine Menge global eindeutiger IP-Adressen zugewiesen. Wenn ein Rechner innerhalb des Intranets mit einem Rechner im Internet kommunizieren möchte, dann ersetzt die NAT-Box die Absenderadresse im IP-Packet durch eine der global eindeutigen Adressen. Wenn Packete von aussen an eine dieser Adressen gesendet werden, dann ersetzt die NAT-Box diese im Packet durch die entsprechende lokale Adresse. Somit benötigt die NAT-Box so viele global eindeutige IP-Adressen, wie lokale Rechner gleichzeitig mit dem Internet kommunizieren können sollen. Die NAT-Box führt dazu eine Tabelle mit den an lokale Rechner dynamisch zugewiesenen globalen IP-Adressen.

Der wesentliche Nachteil von NAT ist der Verstoß gegen die allgemeine Annahme in IP-basierten Protokollen, dass alle Rechner über global eindeutige Adressen erreichbar sind. Wegen dieser Annahme werden oft IP-Adressen auch in Protokollen höherer Schichten übertragen. Dadurch bedingt muss eine NAT-Box auch diese Protokolle verstehen, um die Adressen in den entsprechenden Header-Einträgen dieser Protokolle ebenfalls zu ändern. Dadurch wird die Einführung neuer Protokolle via NAT behindert.

Trotz dieses Nachteils hat die Existenz von NAT den Zwang zur Einführung eines Nachfolger-Protokolls für IP in der derzeit verwendeten Version 4 wesentlich reduziert. Das im Folgenden beschriebene IP Version 6 ist die technologisch bessere Lösung zur Überwindung der Adressknappheit von IP Version 4.

IP Version 6

- Durch das unerwartet (?) starke Wachstum des Internet sind im gegenwärtig verwendeten IP (IPv4) kaum noch freie Adressen verfügbar.
- Da die Vergrößerung des Adreßraums eine Veränderung des IP–Headers verlangt, muß deshalb eine neue Version von IP defniert werden.
  - Diese ist heute als IPv6 bekannt, da Version 5 bereits für ein experimentelles Echtzeit-Datenstrom Protokoll vergeben war.
- IPv6 wird auch manchmal als IPng (*IP next generation*) bezeichnet.
- Da durch die erweiterten Adressen ohnehin ein neues IP–Protokoll eingeführt werden muß, wurden dabei auch weitere Veränderungen vorgenommen.
  - Die Umstellung von IPv4 auf IPv6 bedeutet schließlich die Änderung der Netzwerk-Software in jedem Rechner und jedem Router im Internet.
- Da das Internet sehr groß und ohne zentrale Kontrollinstanz ist, kann die Umstellung nur nach und nach geschehen.
  - Deshalb muß die Koexistenz beider Protokoll–Varianten über Jahre hinweg vorgesehen sein.
- Bislang existieren nur wenige (experimentelle) Netzwerk–Verbindungen, auf denen IPv6 eingesetzt wird.
  - Mit einer weiten Verbreitung ist in den nächsten Jahren zu rechnen.

Haupt-Merkmale

- 128-Bit-Adressen
- klassenloses Adressieren/Routing (vergleichbar mit CIDR)
- Multicast
- Anycast: Routing zu *einer* (beliebigen) Adresse aus einer Gruppe
- Echtzeitservice (Quality of Service)
- Authentisierung und Sicherheit
- Auto–Konfiguration
- Ende–zu–Ende–Fragmentierung

Adressen-Präfixe für IPv6

Prefix	Use
0000 0000	Reserved (including IPv4)
0000 001	OSI NSAP addresses
0000 010	Novell Netware IPX addresses
001	Aggregatable Global Unicast Addresses
1111 1110 10	Link Local Use Addresses
1111 1110 11	Site Local Use Addresses
1111 1111	Multicast Addresses

- Adressierung für:
  - andere Protokolle (IPv4, OSI, Novell)
  - LAN–lokale Knoten (für Auto–Konfiguration)
  - Vom Internet separate Netze
  - Multicast

Beachte: In IPv6 sind noch nicht alle Details fixiert. Deshalb können sich auch die bisher definierten Adressen-Präfixe noch ändern.

Adressen

- Notation x:x:x:x:x:x:x:x (x = 16-bit Hex-Zahl)
  - benachbarte 0-Blöcke können zusammengefasst werden:
    - `47CD:0000:0000:0000:0000:0000:A456:0124`
    - `47CD::A456:0124`
- Darstellung von IPv4–Adressen
  - IPv4-compatible IPv6 address: ::128.42.1.87
  - IPv4-mapped IPv6 address: ::00FF:128.42.1.87
- Aggregatable Global Unicast Addresses

Diese Adressen sind analog zu CIDR spezifiziert: Ein Provider bekommt einen kurzen Adress-Präfix und gibt Teilbereiche davon (mit längerem Präfix) an seine Kunden weiter. Das Problem dabei ist das für einen Kunden nur unter sehr hohem Aufwand mögliche Wechseln des Providers.

- Die wichtigste (und einzig vollständig spezifizierte) Adressierung ist Registry-basiert.

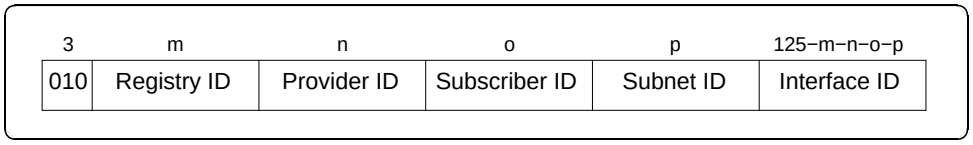

Abb. IPA–1 Registry-basierte IPv6-Adresse

- Registry
  - Verzeichnis von Providern
  - Derzeit sind Registries vergesehen für Nord-Amerika, Europa und Asien
  - $m = 5$, erlaubt insgesamt 32 Registries weltweit
  - n, o, p kann jedes Registry selbst festlegen
- Provider
  - Z.B. große Telefongesellschaften oder Netzbetreiber
    - Entspricht *Transit Autononous System*
- Subscriber
  - Entspricht *Stub* bzw. *Multihomed Autononous System*

Adreßraum

- IPv4 definiert ca. 4 Milliarden Adressen
- IPv6 definiert ca. 3×10^{38} Adressen

Tanenbaum:

"While it was not the intention to give every molecule on the surface of the earth its own IP address, we are not that far off."

- Selbst bei pessimistischer Prognose der Ausnutzung der Adreßräume (analog zur Ausnutzung von Telefonnummern-Kontingenten) kann davon ausgegangen werden, daß mit IPv6 mehr als 1000 IP-Adressen pro Quadratmeter Erdoberfläche (Land und Wasser) zur Verfügung stehen.

Peterson/Davie:

"Based on the most pessimistic estimates of efficiency . . . , the IPv6 address space is predicted to provide over 1500 addresses per square foot of the earth's surface, which certainly seems like it should serve us well even when toasters on Venus have IP addresses."

Header

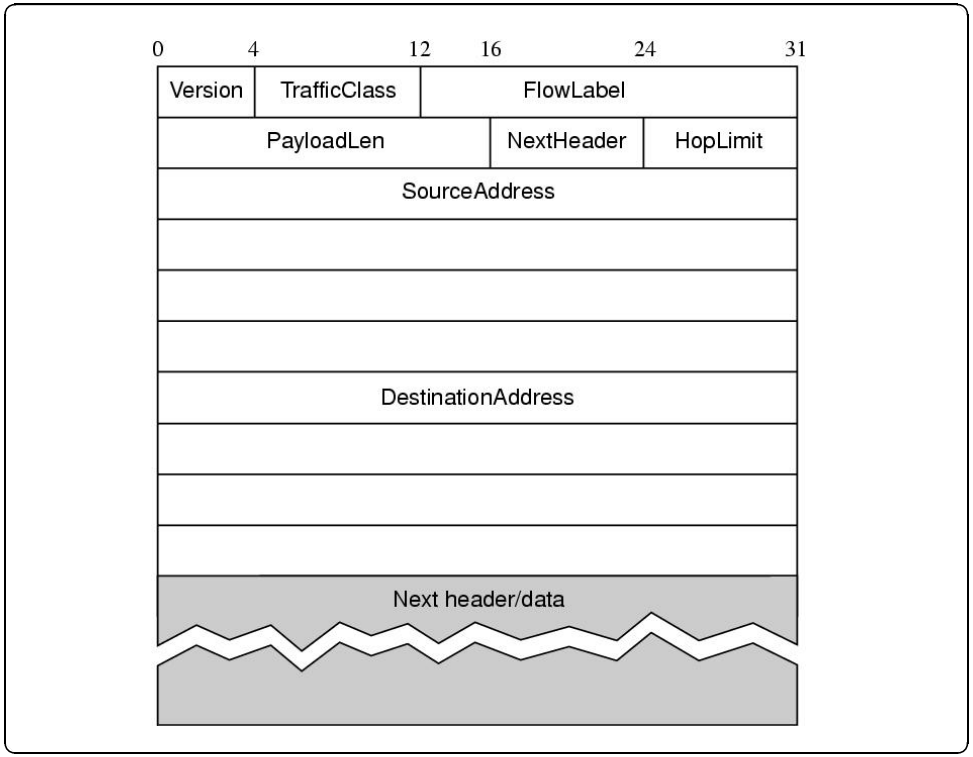

Abb. IPA–2 Der IPv6 Basis–Header

- *Version* = 6
- *Priority,* und *Flow Label,* für Quality of Service – noch nicht voll spezifiziert.
- *Payload Length,* Anzahl Bytes hinter dem 40-Byte Header
- *Hop Limit,* ersetzt IPv4's TTL-Feld
- *Next Header,* Typ des folgenden Headers

  Die einzelnen Header bilden eine verkettete Liste.

  Nur die gerade benötigten Header werden übertragen.
  - *Hop-by-Hop Options,*

  - ◇ z.B. Längenangabe für *Jumbograms* ($> 64K$)
- *Routing*,
  - ◇ Source-Routing Information
- *Fragmentation*
  - ◇ Im Gegensatz zu IPv4 wird Fragmentierung vom Source-Rechner (und nicht mehr von Routern) durchgeführt.
  - ◇ Wenn eine Strecke ein Paket nicht unfragmentiert übertragen kann, wird das Paket weggeworfen und eine ICMP-Fehlermeldung zurückgeschickt.
- *Authentication*,
  - ◇ Identifikation des Absenders
- *Encrypted security payload*,
  - ◇ Information über Verschlüsselung des Inhalts
- *Destination Options*,
  - ◇ Noch nicht spezifiziert
- *Protocol Header*,
  - ◇ Z.B. für TCP, immer zuletzt

Fragmentation Extension Header

Als Beispiel für einen IPv6 Extension Header sei hier der Fragmentation Header gezeigt. Dieser Extension Header wird nur dann in ein Packet eingefügt, wenn wirklich Fragmentierung stattfindet. Ausser den Feldern mit der eigentlichen Fragmentierungs-Information (analog zu den IPv4 Header-Feldern) findet sich vor allem das *NextHeader* Feld am Anfang.

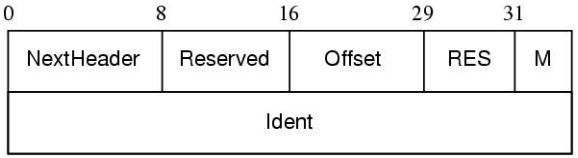

Abb. IPA–3 Der IPv6 Extension–Header für Fragmentation

Auto–Konfiguration

- Problem: Automatische Zuweisung von IP–Adressen bei Installation von Rechnern.
- *Stateful Approach*:
  - Einsatz eines Konfigurations–Servers
- *Stateless Approach*:
  - Selbständige Konfiguration ohne spezielle Server
  - In IPv6 spezifiziert
- Stateless Auto–Konfiguration:

  1. Bestimmung einer Interface-ID, die im Segment eindeutig ist, an dem der Host angeschlossen ist.
  - Z.B. 48-Bit "Ethernet" Adresse

  2. Bestimmung des korrekten Präfix des Segments
  - In einfachen LANs (ohne Router) genügt der *Link Local Use*–Präfix:
    - ◇ 1111 1110 10::Ethernet-Adresse
  - In Router-basierten Netzen versenden die Router regelmäßig die Segment-Adresse per Broadcast.

Advanced Routing

Der Routing Extension Header von IPv6 ermöglicht ein dem Source Routing ähnliches Vorgehen. Dabei kann eine Liste von Knoten oder Areas (z.B. Provider) angegeben werden, die auf dem Weg zum Empfänger überquert werden müssen.

Eine Besonderheit dabei ist die *Anycast* Adressierung für zu passierende Areas; ein beliebiger ("nächster") Knoten einer Area wird dabei ausgewählt. Dies erleichtert die Spezifikation des Weges, weil somit keine Information über die Provider-Netzwerke beim Absender vorhanden sein muss.

Transition von IPv4 zu IPv6

- Dual–Stack Operation
  - Host implementiert beide Protokoll-Stacks.
    - ◇ Unterscheidung durch Versions-Feld im Header
  - Host hat dann auch zwei Adressen, wobei die IPv6-Adresse eine "IPv4-mapped IPv6" Adresse sein kann.
- Tunneling
  - Zur Übertragung von IPv6-Paketen über Strecken, die nur IPv4 implementieren, können IPv6-Pakete in IPv4-Pakete eingepackt werden.
  - Router an den beiden IPv4–Endpunkten packen ein bzw. aus.

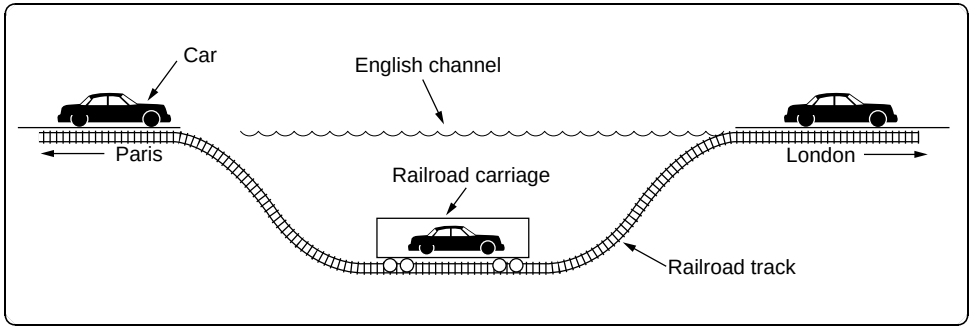

Abb. IPA–4 Metapher für Tunneling von Paketen eines Protokolls durch Kapselung in Paketen eines anderen Protokolls

IP Multicast

Im Folgenden behandeln wir Multicasting in IP-basierten Netzen, vornehmlich dem Internet. Zwischen mehreren Netzen wird das Multicasting in Software realisiert; zu den Endknoten wird (soweit vorhanden) die jeweilige LAN-Technologie zum Multicasting eingesetzt. IP-basiertes Multicasting geschieht meist durch Erweiterung der Routing-Verfahren; davon leitet sich dann auch der jeweilige Name des Multicasting-Verfahrens ab.

Generell werden Multicast-Packete an sogenannte Multicast-Gruppen zugestellt (vergleiche IP-Adressen der Klasse D). Die Zugehörigkeit von Rechnern zu Gruppen wird über das Internet Group Management Protocol (IGMP) geregelt, das hier jedoch nicht diskutiert wird. Prinzipiell jedoch teilt ein Rechner seinem lokalen Router mit, wenn er Packete einer bestimmten Multicast-Gruppe empfangen möchte. Es ist dann die Aufgabe der Router, untereinander für die Weiterleitung von Multicast-Packeten zu sorgen.

Link-State Multicast

Link-State Multicast basiert auf Link-State Routing. Dabei beobachtet jeder Router den Status seiner lokalen Links. Durch den Nachrichten-Austausch bekommt jeder Router die Information über alle Links im Netzwerk und kann dadurch kürzeste Pfade zu den jeweiligen Empfängern von Packeten berechnen.

Beim Link-State Multicast enthalten die zwischen den Routern ausgetauschten Packete zusätzlich Listen der zu empfangenden Multicast Gruppen. Somit können die Router die kürzesten Pfade von einem Sender zur Multicast-Gruppe berechnen. Zur Einsparung von Speicherplatz (in den Routern) werden nur die kürzesten Pfade gerade aktiver Multicast Gruppen gespeichert.

Abb. IPA–5 zeigt ein Beispiel für eine im Internet verteilte Multicast Gruppe. Abb. IPA–6 zeigt die jeweiligen kürzesten Pfade für drei verschiedene Sender. Die Router müssen somit die kürzesten Pfade für alle aktiven Paare von Sendern und Multicast Gruppen speichern.

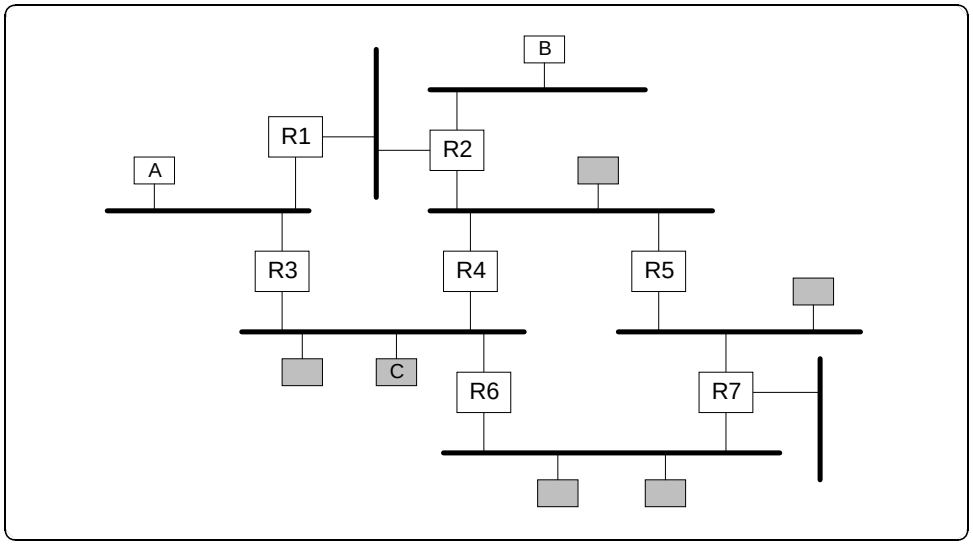

Abb. IPA–5 Beispiel für Internet Multicast
grau schattiert sind die Mitglieder der Multicast-Gruppe

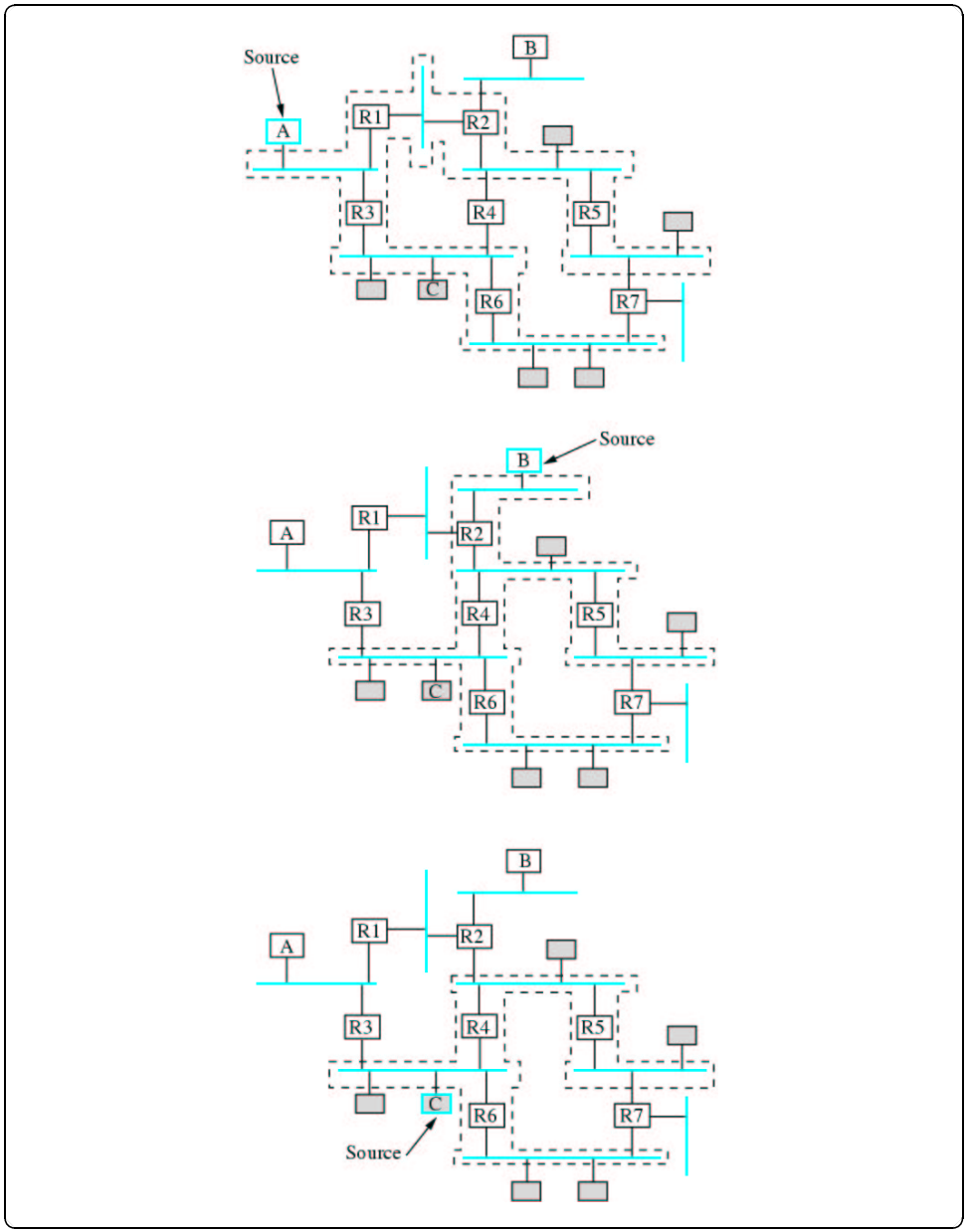

Abb. IPA–6 Kürzeste Pfade für Internet Multicast bei verschiedenen Sendern

Distance-Vector Multicast

Distance-Vector Multicast basiert auf Distance-Vector Routing. Bei diesem Verfahren hat jeder Router eine Tabelle mit Einträgen <Ziel, Kosten, NextHop>. Die Router tauschen untereinander Nachrichten mit ihren Distanz-Vektor Tabellen aus (<Ziel, Kosten>).

Distance-Vector Multicast verwendet als Forwarding Mechanismus das Reverse Path Broadcast Verfahren, das bei hohem Multicast-Verkehrsaufkommen zu Reverse Path Multicast optimiert wird.

- Reverse Path Broadcast (RPB)
  - Jeder Router weiß bereits, daß der kürzeste Pfad zum Ziel S über Router N führt.
  - Wenn ein Multicast-Paket von S erhalten wird, schicke es weiter an alle abgehenden Verbindungen (außer an die eine, auf der das Paket ankommt), aber nur wenn das Paket von N ankommt.
  - Lösche die doppelten Broadcast-Pakete, indem nur dem "Eltern-Router" für LAN (relativ zu S) erlaubt wird, zu senden.
    - ◇ Eltern-Router: kürzester Pfad zu S (lerne aus Distanz-Vektor)
    - ◇ bei Gleichheit: kleinste Adresse

Reverse Path-Multicast (RPM)

- Ziel: Entferne Netzwerke, die keine Hosts in Gruppe G besitzen
- Schritt 1: Ermittlung, welches LAN ein *Blatt* ohne Mitglieder in G ist
  - Blatt, wenn Eltern-Router der einzige Router auf dem LAN ist
  - stelle fest, ob irgendwelche Hosts Mitglieder von G sind
- Schritt 2: Propagiere die "keine Mitglieder von G hier vorhanden" Information
  - erweitere (⟨`Ziel, Kosten`⟩ Update-Nachricht um die Menge der Gruppen, an denen dieses Netzwerk interessiert ist, Multicast-Pakete zu empfangen
  - dies passiert nur, wenn Multicast-Adresse aktiv wird

Protocol Independent Multicast (PIM)

Bei den bisher vorgestellten Multicast Verfahren bekommen alle Router solange Multicast Packete einer Gruppe, bis sie jeweils eine Gruppe aktiv "abbestellen". Dieses Verhalten skaliert sehr schlechtt, weil als Default-Verhalten zu viele Packete gesendet werden.

Das PIM Verfahren wurde vor allem durch seinen *Sparse Mode* (SM) bekannt. Hier wird ein Multicast-Baum explizit aufgebaut; alle unbeteiligten Router werden nicht mit Multicast Packeten belastet.

Bei PIM gibt es einen speziellen *Rendezvous Point* (RP), der allen Teilnehmern im Voraus bekannt sein muss. Router mit lokal angeschlossenen Multicast-Empfängern sender*join* oder *prune* Nachrichten, um Teil des Multicast-Baums zu werden oder diesen wieder zu verlassen. Ein *Shared Tree* (gemeinsamer Baum) mit RP als Wurzel wird dementsprechend aufgebaut. Das Senden von Multicast-Packeten erfolgt zuerst per Tunneling zum RP und von dort aus über den Shared Tree. Bei zu hohem Verkehrsaufkommen kann der Shared Tree auch dynamisch umkonfiguriert werden.

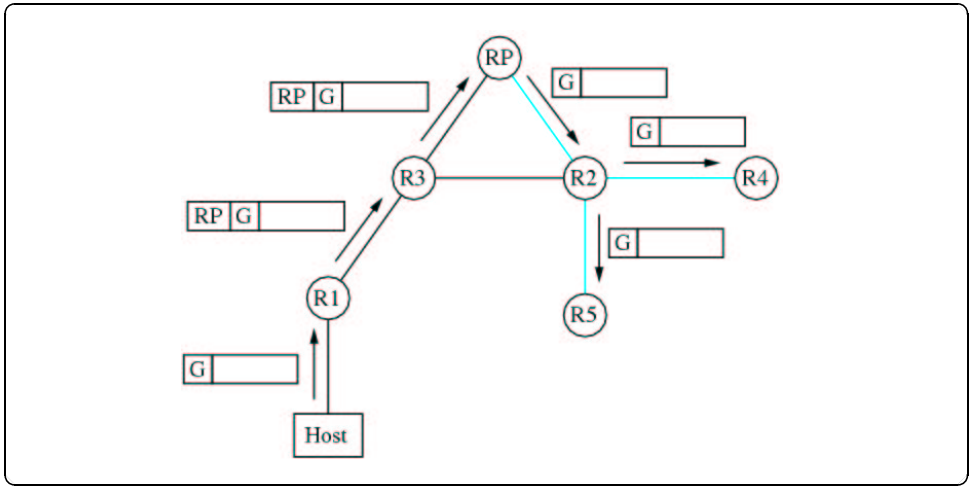

Abb. IPA–7 PIM: Multicast über einen *Shared Tree*

Multicast Backbone (MBone)

MBone ermöglicht Weitverkehrs-Multicast über eine logische Schicht oberhalb des Internet. Am MBone partizipierende Router verwenden Klasse-D Adressen und Tunneling zur Überbrückung von Verbindungen ohne MBone Funktionalität. Es wird DVMRP eingesetzt, das *Distance Vector Multicast Routing Protocol*. MBone ist ein sogenanntes *Testbed*, eine Installation zur Erprobung von Multicast Anwendungen. Ein wichtiges Beispiel hierfür sind die Videokonferenz-Übertragungen der IETF-Meetings.