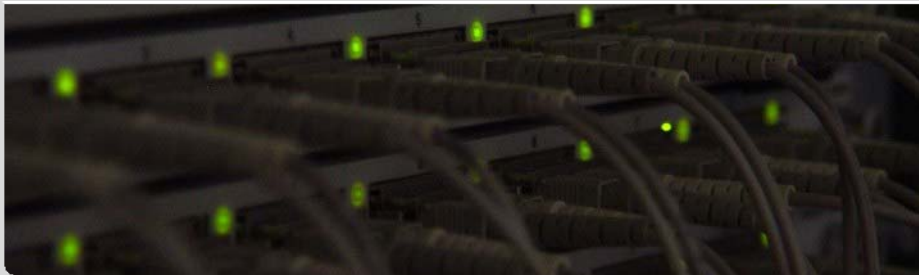


# Next Generation Internet

## 6. Multicast-Routing

INSTITUT FÜR TELEMATIK



KIT – Universität des Landes Baden-Württemberg und  
nationales Forschungszentrum in der Helmholtz-Gemeinschaft

[www.kit.edu](http://www.kit.edu)

## Überblick Kapitel 6

### I. Einführung

1. Einführung

### II. Internet-Architektur

2. Internet-Architektur
3. NAT & IPv6
4. Dienstgüte

### III. Multicast

5. Grundlagen
6. Multicast Routing
7. Multicast Transport

- 6.1 Einleitung
- 6.2 Gruppenverwaltung
- 6.3 Reichweite und Multicast-Adressen
- 6.4 Allokation von Multicast-Adressen
- 6.5 Konzepte zum Multicast-Routing
- 6.6 Multicast-Routingprotokolle im Internet

### IV. Flexible Dienste und Services

8. Neuere Transportprotokolle
9. Aktive Netze
10. Peer-to-Peer

2

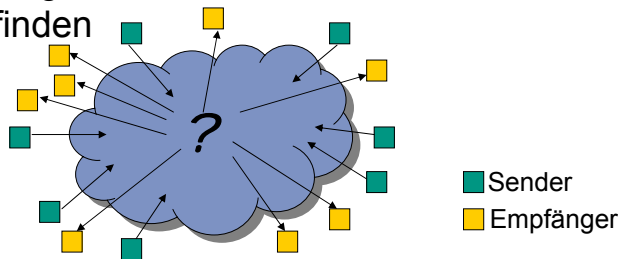
Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
<http://tm.kit.edu/>

## 6.1 Zielsetzung Multicast-Routing

- Multicast-Pakete müssen effizient zugestellt werden
  - Replikation so spät wie möglich anhand Gruppenzugehörigkeit
  - Gruppenadresse im Paket
- Paket muss Weg von den Sendern zu den Empfängern finden



3

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
<http://tm.kit.edu/>

- Gruppenadresse ist eine logische Zieladresse und umfasst/bezeichnet die Menge aller Empfänger
- Problem ist hier u.a., dass die Gruppenzugehörigkeit sich dynamisch ändern kann.

## Zweigeteilter Mechanismus

- Router im Subnetz kennen Zugehörigkeit der Endsysteme zur Multicastgruppe  
→ [Protokolle zur Gruppenverwaltung](#)
- Router koordinieren sich untereinander, um Pakete für eine bestimmte Gruppe von einem oder mehreren Sendern zu erhalten  
→ [Multicast-Routingprotokolle](#)

4

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
<http://tm.kit.edu/>

## 6.2 Gruppenverwaltung



### ■ Problem

- Woher weiß ein Router, dass er Multicast-Dateneinheiten an die angeschlossenen Subnetze bzw. die darin lokalisierten Systeme weiterleiten muss?

### ■ Lösung

- Multicast-Empfänger informieren „ihren“ Multicast-Router über ihre Gruppenmitgliedschaft(en). Im Internet werden hierzu eingesetzt:
  - Für IPv4: **IGMP** (Internet Group Management Protocol, aktuell: **IGMPv3**)
  - Für IPv6: **Multicast Listener Discovery (MLD)** für IPv6 (integriert in ICMPv6, aktuell: **MLDv2**)

5

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
<http://tm.kit.edu/>

## Gruppenverwaltung: Allg. Ablauf



### ■ Genereller Ablauf (IPv6)

- Ändert sich die Gruppenmitgliedschaft eines Systems, wird eine „**Membership-Report**“-Nachricht mit der entsprechenden Zustandsänderung geschickt
- Multicast-Router senden außerdem periodisch so genannte „**Membership-Query**“-Dateneinheiten an die Multicast-Adresse „**all nodes**“ (ff02::1)
- Jeder Multicast-Empfänger im Subnetz sendet, nach einer zufälligen Wartezeit, als Antwort eine oder mehrere „**Membership-Report**“-Dateneinheiten an „**all MLDv2-capable routers**“ (ff02::16), in welchen die Adressen der gewünschten Multicast-Gruppen enthalten sind

6

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



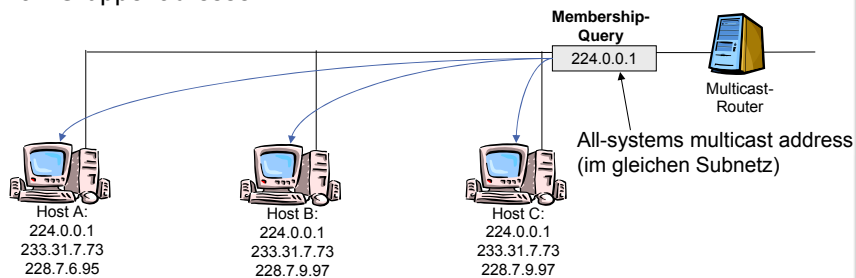
Institut für Telematik, Fakultät für Informatik  
<http://tm.kit.edu/>

## IGMP: Ablauf (1)



### ■ Multicast-Router

- tritt Multicast-Adresse 224.0.0.22 bei, um Membership Report zu empfangen
- Periodische Anfrage (Variante: „General-Query“) an „all-systems“ (224.0.0.1)
  - Auffrischen der Zustandsinformation
- Sendet Group-Specific- und Group-and-Source-Specific-Queries gezielt an Gruppenadresse



7

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
http://tm.kit.edu/

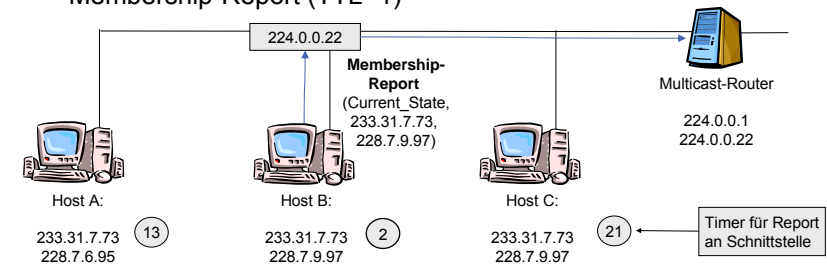
- Empfangene Membership-Reports, die Zustandsänderungen anzeigen, werden sofort bearbeitet
- Periodische Anfrage General-Query wird normalerweise alle 125s wiederholt.
- Jedes Multicast-fähige System (Host/Router) ist Mitglied der Gruppe 224.0.0.1
- Die Anfrage wird mit einer TTL (Time-to-Live) von Eins gesendet (Weshalb?)

## IGMP: Ablauf (2)



### ■ Multicast-Hosts

- Beitritt zu/Austritt aus einer Multicast-Gruppe: sofortiges Senden eines Membership-Reports
- Jedes System startet nach Empfang eines Membership-Queries einen Timer (bei General-Query je Schnittstelle, bei den anderen Varianten zusätzlich je Gruppe bzw. Quelladresse)
- Läuft einer der Timer ab, so antwortet dieser Rechner mit einem Membership-Report (TTL=1)



8

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
http://tm.kit.edu/

- Membership Report wird im Allgemeinen einmal nach 1s wiederholt; die Anzahl der Wiederholungen wird durch eine sogenannte „Robustheits-Variable“ festgelegt, die standardmäßig auf 2 gesetzt wird.
- Bei IGMPv2 werden weitere Membership Reports für die gleiche Gruppe unterdrückt. Für den Router ist im Wesentlichen wichtig zu wissen, dass mindestens ein Endsystem an der entsprechenden Multicast-Gruppe teilnimmt.
- Problem
  - Subnetz verfügt möglicherweise über mehrere Multicast-Router
  - Lösung: Multicast-Router kann bei IGMP zwei Rollen einnehmen: Querier oder Non-Querier
  - Querier
    - In jedem Subnetz existiert genau ein Querier (niedrigste IP-Adresse)
    - Verantwortlich für das periodische Nachfragen im Subnetz
    - Falls General-Query von niedrigerer IP-Adresse kommt, hört er auf General-Queries zu senden und wird anschließend zum
  - Non-Querier
    - Versendet keine periodischen Nachfragen
    - Setzt Other-Querier-Present-Timer
    - Setzt Timer zurück sobald General-Query von niedrigerer IP Adresse empfangen wird
    - Läuft Timer ab, wird er zum Querier und versendet er General-Queries

## IGMPv3



- Unterstützung für **Source-Specific Multicast (SSM)**: Erweiterung um „**Source Filtering**“
- Einige Änderungen gegenüber IGMPv2:
  - Zustand wird für **Gruppe und Liste von Quellen** gehalten (nicht nur für Gruppe)
  - Programmierschnittstelle erlaubt nun die **Angabe von Quellen-Listen**
  - **Hosts unterdrücken Membership-Reports nicht mehr** (vereinfachte Implementierung und explizites Verfolgen der Gruppen-Mitgliedschaft)
  - Reports können mehrere Gruppeneinträge enthalten
  - Reports werden an 224.0.0.22 geschickt (erlaubt einfacheres „Snooping“ durch Schicht-2-Switches)

9

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
<http://tm.kit.edu/>

- IGMPv3 ist in RFC 3376 definiert und rückwärtskompatibel zu IGMP Versionen 1 und 2
- „Source Filtering“ bietet die Möglichkeit nur Interesse für Multicast-Pakete mit einer bestimmten Quelladresse (also eines bestimmten Senders) auszudrücken
- Weitere Änderungen gegenüber IGMPv2:
  - Hosts wiederholen Zustandsänderungsnachrichten
  - Querier fügt Robustheits-Variable und Query-Intervall zu Queries hinzu (erlaubt Synchronisation der Non-Queriers)

## 6.3 Reichweite und Multicast-Adressen



- **Reichweite**
  - Definiert den Bereich, in dem die Multicast-Dateneinheit weitergereicht wird
- **Vorteil einer begrenzten Reichweite**
  - Begrenzung von gefluteten Netzbereichen
    - z.B. beim Routingprotokoll DVMRP
  - Mehrfachnutzung von Multicast-Adressen in verschiedenen Netzbereichen
  - Privatsphäre
- **TTL-Scoping**: Begrenzung anhand der TTL-Werte
  - Verwendung von Schwellenwerten
  - Bereiche: Subnetz (1), Domäne, Region, Kontinent usw.

10

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
<http://tm.kit.edu/>

- Privatsphäre kann in der Regel nur bedingt erreicht werden.
- TTL-Scoping: Falls TTL-Wert kleiner als Schwellenwert, wird die Dateneinheit verworfen

Schwellenwert	Reichweite	
	0	Begrenzung auf einen Knoten
	1	Begrenzung auf ein Subnetz
	32	Begrenzung auf eine Domäne
In Europa:	48	Begrenzung auf ein Land
	64	Begrenzung auf eine Region
	128	Begrenzung auf ein Kontinent
	255	unbegrenzt

## Reichweite von Multicast-Gruppen



### ■ Administrative Bereiche für IPv4 [RFC2365]

- Die Multicast-Adresse gibt die Reichweite an und muss deshalb entsprechend gewählt werden

Adressenbereich	Bereichsbezeichnung
224.0.0.0–224.0.0.255	Link Local
224.0.1.0–238.255.255.255	Global
239.0.0.0–239.255.255.255	Administratively Scoped:
239.192.0.0–239.251.255.255	Organization-Local
239.255.0.0–239.255.255.255	Site-Local

### ■ Für IPv6 in Multicast-Adressstruktur mit definiert: [RFC4291]

- 4 Scope-Bits: 16 mögliche Gültigkeitsbereiche/Reichweite

11

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
<http://tm.kit.edu/>

- Lokaler Bereich (Site-Local), z.B. innerhalb eines Firmennetzes
- Organisatorischer Bereich (Organization-Local)
  - Einer Organisation wird ein Adressbereich zugewiesen.
  - Beispiel: Breitband-Wissenschaftsnetz des Deutschen Forschungsnetzes (DFN)
- Feinere Granularität bei der Berechnung der Reichweite
- Administrative Bereiche müssen bekannt sein
- Site-Local-Bereiche dürfen sich nicht überlappen

## 6.4 Allokation von Multicast-Adressen



### ■ Bisher kein Zuteilungsverfahren für Multicast-Adressen im Internet

- für Source-Specific-Multicast auch nicht notwendig
- für traditionelles Any-Source-Multicast: Kollisionen möglich
- Wahrscheinlichkeit einer Kollision nimmt mit steigender Nutzung zu
- Wie wird bisher vorgegangen?

### ■ Multicast Address Allocation Architecture: Vorschlag zur Zuteilung von Multicast-Adressen (MALLOC Working Group der IETF)

- Ansatz zu komplex → keine Umsetzung i.d. Praxis
  - Client/Server Architektur, hierarchisch

12

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
<http://tm.kit.edu/>

- Ziele
  - Geringe Kollisionswahrscheinlichkeit
  - Aggregation von Adress-Bereichen
- Basiert auf administrativen Bereichen
- Reservierungsarten
  - Statisch: feste Zuteilung (z. B. Session Announcement Protocol SAP)
  - Bereichs-relativ: reserviert für Infrastruktur-Protokolle, die eine Adresse in allen administrativen Bereichen benötigen
  - Dynamisch:
    - Zuteilung erfolgt auf Anfrage
    - begrenzte Reichweite, gleiche Adresse kann in verschiedenen administrativen Bereichen verwendet werden
- Multicast Address Allocation Server (MAAS) verwalten Adressraum dezentral
- Multicast Address Dynamic Client Allocation Protocol (MADCAP)
  - Client-Server-Protokoll mit folgenden Aufgaben
    - Auffinden eines lokalen MAAS; Anfordern, Verlängern und Freigabe von Adressen; Konfigurationsinformationen
- Multicast Address Allocation Protocol (AAP)
  - Intra-Domain-Server-Protokoll mit folgenden Aufgaben
    - Garantie der Eindeutigkeit von Adressen innerhalb einer Domäne
    - Bekanntgabe vergebener Adressen; Zuweisung und Reservierung von Adressen; Bekanntgabe von Adressbereichen
  - Seit Mai 2001 nicht mehr weiter verfolgt, da Kooperation zwischen mehreren MAAS in einer Domäne auch proprietär erfolgen kann. Für Zero-Configuration Networks wurde das Zeroconf Multicast Address Allocation Protocol (ZMAAP) entwickelt.
- Multicast Address-Set Claim (MASC)
  - Inter-Domain-Server-Protokoll
  - Hierarchische Organisation von MASC-Routern (Eltern, Kinder, Geschwister, interne MASC-Router)
  - Aufgabe: Zuweisung von Adress-Bereichen zu Domänen

## Wdh.: IPv6 Multicast-Adressen



### Struktur [RFC4291]

11111111	Flags	Scope	Group ID
8	4	4	112 Bits

#### Flags

0 R P T

##### Transient-Bit

- 0 = permanente, „wohlbekannte“ (von der IANA vergebene) Gruppenadresse
- 1 = dynamisch vergebene, „transiente“ Gruppenadresse

##### P-Flag für Unicast-Prefix-based IPv6 Multicast [RFC3306]

##### R-Flag für eingebettete Rendezvous-Point-Adressen [RFC3956]

#### Scope

- 6 verschiedene Bereiche

13

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
http://tm.kit.edu/

## IPv6 Unicast Prefix-based Addresses



### Unicast Prefix-based IPv6 Multicast Addresses [RFC3306]

11111111	Flags	Scope	Reserved	Plen	Network Prefix	Group ID
8	4	4	8	8	64	32 Bits

0 0 P T

- P=1: Multicast-Adresse basierend auf Netzpräfix

- Plen: Präfixlänge des Subnetzes (in Bit)

- Network Prefix: Unicast-Subnetz der ausgebenen Domäne

#### Vorteil

- keine Adressenkollisionen mehr möglich
- Vereinfachung: Multicast-Adressenallokationsprotokolle können wegfallen (AAP und MASC)

14

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
http://tm.kit.edu/

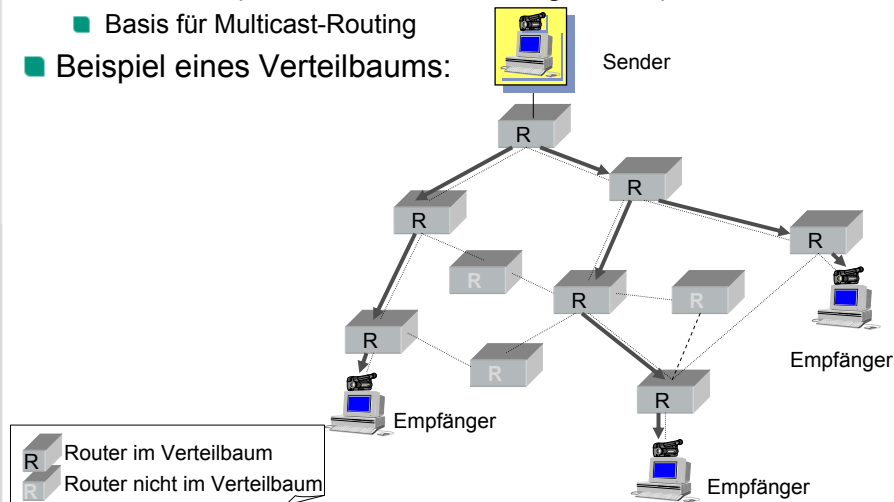
## 6.5 Konzepte zum Multicast-Routing



### ■ Verteilbaum (auch Multicast-Baum genannt)

- Basis für Multicast-Routing

### ■ Beispiel eines Verteilbaums:



15

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
<http://tm.kit.edu/>

## Multicast Routing-Problem



### ■ Problem

- Paket von Sender muss zu den Empfängern finden  
→ Ziel: Aufbau eines effizienten Verteilbaums
- Empfänger müssen Weg zum Sender bzw. zum Verteilbaum finden

### ■ Zwei grundlegende Ansätze

#### ■ Fluten und Zurückschneiden (Flood & Prune)

- Fluten notwendig, weil nicht bekannt ist, wo potentielle Empfänger zu finden sind
- Zurückschneiden: Dort wo keine Empfänger vorhanden sind, wird der Verteilbaum „beschnitten“

#### ■ Bäume mit Rendezvous-Stellen

- Rendezvousstellen dienen dazu, dass
  - Sender zu den Empfängern bzw. zum Verteilbaum finden
  - Empfänger zu den Sendern bzw. zum Verteilbaum finden

16

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
<http://tm.kit.edu/>



## 6.5.1 Fluten und verbesserte Techniken



### ■ Fluten

- Daten werden über alle Anschlüsse eines Routers geflutet, um alle Gruppenmitglieder zu erreichen
- Vorteile:
  - einfache Realisierbarkeit
  - Erreichbarkeit aller Gruppenmitglieder garantiert
- Nachteile:
  - hohe Netzlast
  - Keine Unterstützung geschlossener Gruppen möglich
  - Duplikate von Dateneinheiten möglich

### ■ Verbessertes Fluten

- Keine Weiterleitung, falls Duplikat einer Dateneinheit bereits empfangen wurde
- Vorteil:
  - Vermeidung von Schleifen
- Nachteile:
  - Information über empfangene Daten muss gespeichert werden
  - Löschen dieser Information muss kontrolliert werden

17

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
<http://tm.kit.edu/>

## Spannbäume



### ■ Spannbäume

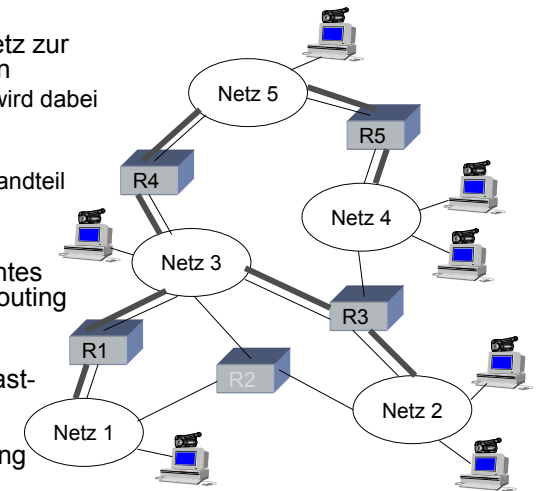
- Etablieren ein Overlay-Netz zur Vermeidung von Schleifen
  - Gruppenmitgliedschaft wird dabei nicht berücksichtigt

### ■ Beispiel

- Router R2 ist nicht Bestandteil des Spannbaums

### ■ Beurteilung

- Fluten stellt sehr ineffizientes Verfahren für Multicast-Routing dar
- Es handelt sich dabei eigentlich um ein Broadcast-Verfahren, weil keinerlei Information über die Gruppenzusammensetzung ausgenutzt wird



18

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
<http://tm.kit.edu/>

## Zielsetzung



### ■ Ziel

- Aufbau eines Verteilbaums unter Berücksichtigung der Gruppenmitgliedschaft
- Multicast-Pakete werden nur an die Empfänger aus der entsprechenden Gruppe weitergeleitet

### ■ Varianten

- Flood & Prune
  - Spannbäume
- Bäume mit Rendezvous-Stellen

19

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
<http://tm.kit.edu/>

## 6.5.2 Flood & Prune



### ■ Vorgehen

- Für jeden Empfänger wird ein Spannbaum etabliert. Optimierungskriterium sind die Verzögerungszeiten.

### ■ RPF-Verfahren (Reverse Path Forwarding)

- RPF-Router merkt sich Datenquelle und Schnittstelle  $S_i$ , über welche die Multicast-Dateneinheit empfangen wurde
- Gehört die Schnittstelle **zum kürzesten Pfad zur Quelle**, so wird die Dateneinheit an alle anderen Schnittstellen weitergeleitet (andernfalls nicht)

### ■ Eigenschaften

- Keine Schleifen möglich
- Mehrfache Routen möglich
- Vorteil
  - Keine gesonderten Multicast-Routingtabellen erforderlich
  - Reduzierter Overhead im Vergleich zu Fluten
- Nachteil
  - Weiterleiten der Daten erfolgt nicht zielorientiert
  - Empfänger kann Daten mehrfach erhalten

20

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
<http://tm.kit.edu/>

## Beispiel zu Reverse Path Forwarding

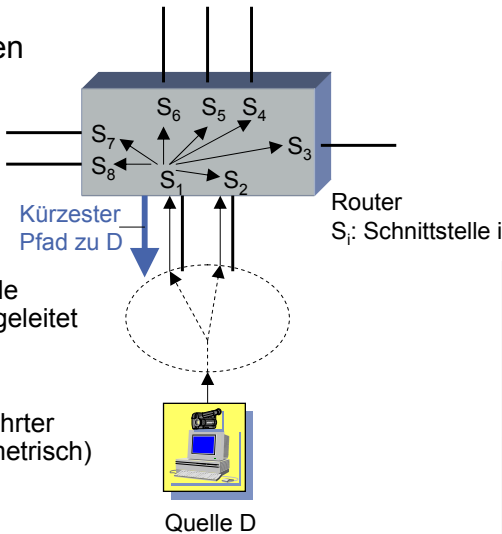


### ■ Endsystem D sendet Daten

- Schnittstelle  $S_1$  liegt auf dem kürzesten Pfad zu D, d.h., Daten werden an alle anderen Schnittstellen weitergeleitet
- Schnittstelle  $S_2$  liegt nicht auf dem kürzesten Pfad, von D über  $S_2$  ankommende Daten werden nicht weitergeleitet

### ■ Problem

- Entscheidung basiert auf kürzestem Pfad in umgekehrter Richtung (Pfade oft asymmetrisch)
- Gruppenstruktur wird nicht berücksichtigt



21

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
http://tm.kit.edu/

- Reverse Path Forwarding verwendet den kürzesten Pfad zur Quelle, d.h. in Richtung, die entgegengesetzt zur Weiterleitungsrichtung ist

## Verbesserung: kürzester Pfad



### ■ Vorgehen

- Zusätzlich zu RPF wird der kürzeste Pfad auch für das Weiterleiten der Daten ausgewertet
- Weiterleitung erfolgt nur noch an solche Router, die in umgekehrter Richtung auf dem kürzesten Pfad liegen

### ■ Beispiel

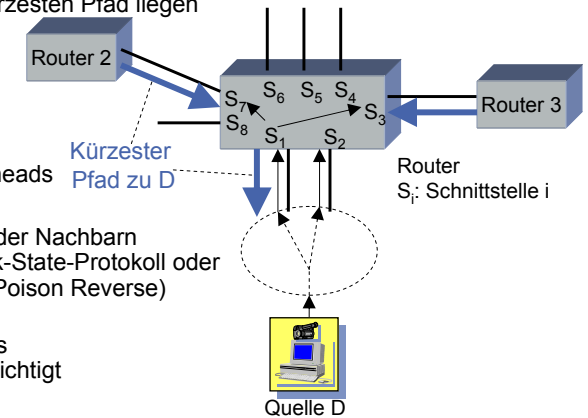
- Weiterleitung an  $S_3$  und  $S_7$

### ■ Vorteil

- Reduktion des Overheads

### ■ Nachteil

- Routing-Information der Nachbarn erforderlich (z.B. Link-State-Protokoll oder Distance Vector mit Poison Reverse)
- Struktur der Gruppe wird beim Aufbau des Baums nicht berücksichtigt



22

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
http://tm.kit.edu/

## Verbesserung: Teilnetze mit Gruppenmitgliedern



### ■ Ziel

- Daten sollen nicht an Schnittstelle (Teilnetz) weitergeleitet werden, in dem sich kein Gruppenmitglied befindet

### ■ Voraussetzung

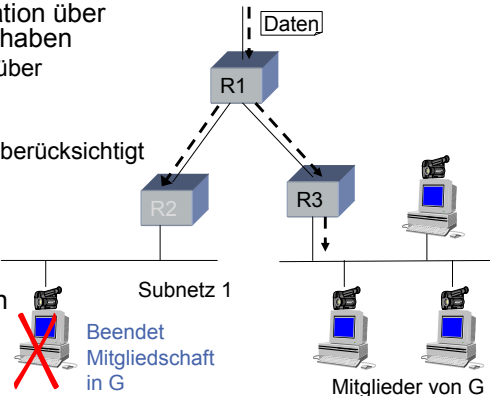
- Router müssen Information über Gruppenmitgliedschaft haben
  - Diese Information ist über IGMP/MLD verfügbar

### ■ Vorteil

- Gruppenstruktur wird berücksichtigt

### ■ Beispiel

- Router R2 besitzt kein Gruppenmitglied von G in Subnetz 1
- Daten für G werden von R2 nicht in Subnetz 1 weitergeleitet



23

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
http://tm.kit.edu/

## Verbesserung: Netzbereiche mit Gruppenmitgliedern



### ■ Ziel

- Daten sollen nicht zu Routern gelangen, hinter denen sich keine Gruppenmitglieder mehr befinden

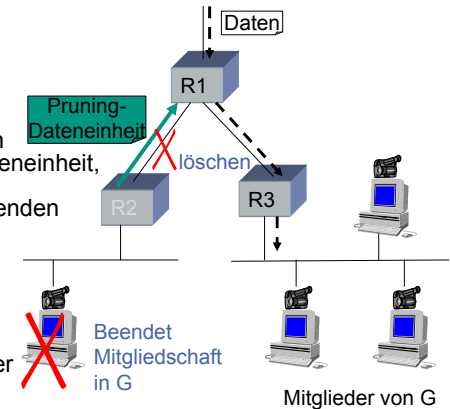
### ■ Vorgehen

#### ■ Flood & Prune

- Daten werden periodisch im Netz geflutet
- Downstream-Router reagieren mit Senden einer **Pruning**-Dateneinheit, falls sie keine aktuellen Gruppenmitglieder zur betreffenden Gruppe kennen (vgl. R2)

#### ■ Expliziter Beitritt (Join)

- Downstream-Router signalisiert Interesse mit Join-Nachricht
  - Erfordert Kenntnis über Sender
    - Rendezvous-Stelle
    - Source-Specific Multicast



24

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
http://tm.kit.edu/

## Verfahren mit Fluten



Verfahren	Broad-/Multicast	Weiterleiten	Dateneinheiten	Netzbereich
Fluten	Broadcast	alle Schnittstellen	alle Dateneinheiten	komplettes Netz
RPF	Broadcast	alle Schnittstellen	Dateneinheiten vom kürzesten Pfad	komplettes Netz
Verbesserung: kürzester Pfad	Broadcast	Schnittstellen auf dem kürzesten Pfad	Dateneinheiten vom kürzesten Pfad	komplettes Netz
Verbesserung: Teilnetze	Multicast	Schnittstellen auf dem kürzesten Pfad	Dateneinheiten vom kürzesten Pfad	nicht in Teilnetze ohne Mitglieder
Verbesserung: Netzbereiche	Multicast	Schnittstellen auf dem kürzesten Pfad	Dateneinheiten vom kürzesten Pfad	nur in Netzbereiche mit Mitgliedern

25

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
http://tm.kit.edu/

## 6.5.3 Bäume mit Rendezvous-Stellen



### ■ Ziel

- Vermeiden des anfänglichen flächendeckenden Flutens

### ■ Vorgehensweise

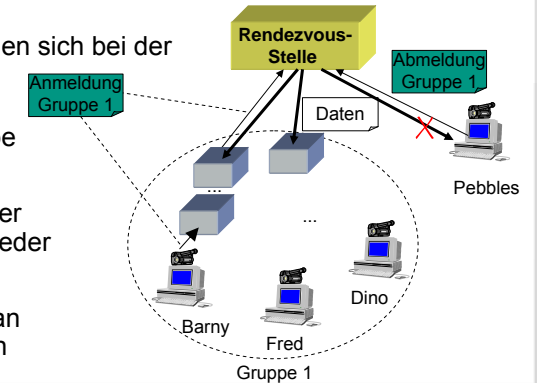
- Für eine Gruppe werden ein oder mehrere **Rendezvous-Stellen** selektiert
- Gruppenmitglieder melden sich bei der Rendezvous-Stelle an
- Etablierung eines Spannbauums pro Gruppe

### ■ Vorteil

- Gezieltes Weiterleiten der Daten an Gruppenmitglieder

### ■ Nachteil

- Verkehrskonzentration an den Rendezvous-Stellen



26

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
http://tm.kit.edu/

- Rendezvous-Stellen heißen englisch Rendezvous Point

## Gegenüberstellung der Basisverfahren



Kriterium	Flood & Prune	Rendezvous-Stellen
Ansatz	inkrementell	inkrementell
Fluten	ja, teilweise eingeschränkt	nein
Zentral/verteilt	verteilt	verteilt
Single-Point-of-Failure	nein	ja
Gruppendichte	hoch	gering
Verkehrskonzentration	nein	ja
Praxis	ja	ja
Aufwand	gering	mittel

27

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
http://tm.kit.edu/

## 6.6 Multicast-Routingprotokolle im Internet



### ■ Ziel

- Vorstellung standardisierter Multicast-Routingprotokolle, die im Internet bzw. in Teilbereichen eingesetzt werden
- Präsentation neuer Ansätze, die sich noch in der Standardisierung befinden

### ■ Grundlage

- Alle vorgestellten Protokolle bauen auf IGMP/MLD für die Gruppenverwaltung auf
- Bei gemeinsamem Medium wird vorausgesetzt, dass ein so genannter Designated-Router über die Mitgliedschaft informiert
- Nur der Designated-Router leitet Multicast-Pakete an Empfänger innerhalb eines gemeinsamen Subnetzes weiter

28

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
http://tm.kit.edu/

## IP Multicast – ASM contra SSM (1)



### ■ Any-Source Multicast (ASM)

- ursprüngliches IP-Multicast-Konzept nach RFC 1112 basiert auf offenen Gruppen (auch Nicht-Gruppenmitglieder dürfen an Gruppe senden)
- Jedes Endsystem kann Pakete an eine Multicast-Gruppe G (Host Group) schicken, Sender muss nicht gleichzeitig Empfänger sein
- Multipeer (m:n-Kommunikation, aber als Multicast im Internet bezeichnet)
  - Unterstützt 1:n-Kommunikation (Multicast) als Spezialfall

### ■ Probleme:

- Adressen-Allokation erforderlich (komplex)
- Fehlende Zugriffskontrolle
  - Empfänger kann Empfang nicht auf bestimmte Quellen einschränken
  - Aktive Quelle kann andere Sender nicht am Senden hindern
- Ineffiziente Behandlung bei wohlbekannten Quellen: Gemeinsam genutzter Verteilbaum unnötig

29

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
<http://tm.kit.edu/>

## IP Multicast – ASM contra SSM (2)



### ■ Source-Specific Multicast (SSM)

- Echtes Multicast (1:n-Kommunikation)
- Kanalkonzept:
  - „Kanal“ (S,G) umfasst jetzt die Quelladresse S des Senders sowie die Gruppenadresse G (IP-Zieladresse) und identifiziert damit eine SSM-Gruppe eindeutig
  - damit  $(S1,G) \neq (S2,G)$
- Ermöglicht zahlreiche Vereinfachungen, u.a.
  - Adressallokation
  - Zugriffskontrolle
  - Vermeidung gemeinsam genutzter Verteilbäume
  - Routing, da Sender einfach gefunden werden kann (s. nächste Folie)

30

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
<http://tm.kit.edu/>

- Für IPv6 Unicast-Prefix-based Multicast Adressen gilt:  
SSM-Gruppenadressen (G) haben folgenden Aufbau: P=1, Plen=0, Network Prefix= 0 → FF3x::/32,  
x=Scope

## SSM vereinfacht Routing



- Da Sender bekannt, können die Empfänger zum Sender finden
- Aufbau des **senderspezifischen Verteilbaums**
  - Empfänger schickt „Join“ in Richtung Sender mittels normalem Routingprotokoll
  - Dort wo Router bereits im Verteilbaum eingegliedert ist, hört das „Join“ auf
- Vorteil
  - Kein Fluten notwendig
- Nachteil
  - Routing-Information wird für Weiterleitung in umgekehrter Richtung verwendet, daher ggf. suboptimale Routen

31

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
<http://tm.kit.edu/>

## Beispiele für Multicast-Routingprotokolle



- Intra-Domain-Routingprotokolle
  - **DVMRP** (Distance Vector Multicast Routing Protocol)
  - **MOSPF** (Multicast Open Shortest Path First)
  - **PIM** (Protocol Independent Multicast)
    - Sparse-Mode (**PIM-SM**)
    - Dense-Mode (**PIM-DM**)
    - Bidirectional PIM (**BIDIR-PIM**)
  - **CBT** (Core-Based-Trees) [RFC2201, RFC2189]
    - Nicht weiter behandelt, da praktisch kaum von Bedeutung
    - Stellt bidirektionale gemeinsame Verteilbäume bereit
    - Probleme
      - Abbildung der Gruppenadresse auf Core (Wurzel des Verteilbaums)
      - gute Lokation des Cores
- Inter-Domain-Routingprotokolle
  - **BGMP** (Border Gateway Multicast Protocol)
  - **Multicast Source Discovery Protocol (MSDP)**

32

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
<http://tm.kit.edu/>



## 6.6.1 Designated-Router

### ■ Problem

- In einem Teilnetz befinden sich mehrere multicastfähige Router. Multicast-Daten können so mehrfach in ein Teilnetz gelangen.

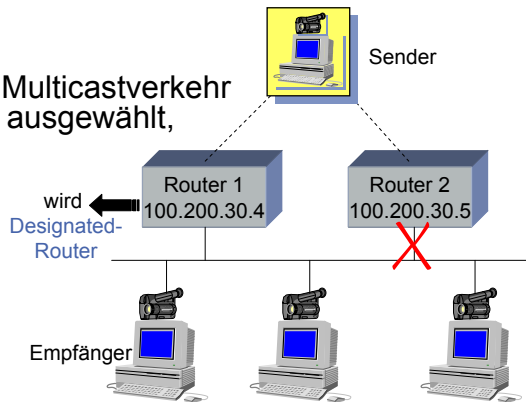
### ■ Verfahren

- Es wird ein für den Multicastverkehr zuständiger Router ausgewählt, der so genannte **Designated-Router**

- Auswahlkriterium: niedrigste IP-Adresse

### ■ Beispiel

- Router R1 wird Designated-Router



## 6.6.2 DVMRP

### ■ Erweiterung der Konzepte von RIP für Multicast-Routing

- Empfängerbasiertes Verfahren
- Berechnet vorangegangenen Übertragungsabschnitt zurück zur Quelle. RIP berechnet Weg hin zur Quelle.

### ■ DVMRP-Router nutzt zwei unabhängige Routingprotokolle

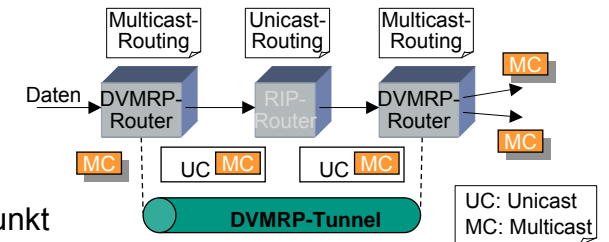
- Protokoll für Unicast-Routing (z.B. RIP, OSPF)
- DVMRP für Multicast-Routing

### ■ Tunneltechnik für den praktischen Einsatz im Internet (ohne Multicast-Unterstützung)

### ■ Beispiel

- Tunnel zwischen DVMRP-Router

### ■ Konzept „virtuelle Schnittstelle“: physikalische Schnittstelle oder Tunnelendpunkt



## Nachbarschaft



### ■ Basis

- DVMRP-Router müssen ihre Nachbarn kennen lernen
- Periodisches Versenden von „Neighbor Probe“-Dateneinheiten
  - TTL ist auf Eins gesetzt
    - Was bewirkt dies?
  - Liste mit allen Nachbarn des Routers

### ■ Routing von Multicast-Daten

- Basiert auf RPF mit Verbesserung: Netzbereiche
- Aufbau der Routingtabelle durch den Austausch von Distanzvektoren zwischen Routern
  - Spezielle Multicast-Distanzvektoren
  - Weg einer Dateneinheit für Unicast und Multicast kann unterschiedlich sein

35

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
<http://tm.kit.edu/>

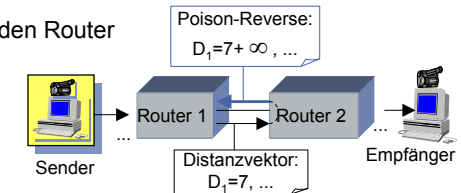
## Dynamische Kontrolle des Multicast-Baums (1)



### ■ Poison-Reverse

- Ziele
  - Dem vorangehenden Router im Baum wird mitgeteilt, dass sich stromabwärts abhängige Router (für eine bestimmte Quelle) befinden
  - Der Router kann eine Liste mit abhängigen, stromabwärts gelegenen Routern je Schnittstelle erstellen
- Ablauf
  - Router bestimmt beste Route bzw. besten Router stromaufwärts in Richtung einer bestimmten Quelle
  - Kennzeichnung der besten Route stromaufwärts durch Erhöhen der Distanz um  $\infty$
  - Reflektieren des Distanzvektors, d.h. er wird an den Router stromaufwärts geschickt
  - Router stromaufwärts trägt den Router stromabwärts in Liste der abhängigen Router ein

### ■ Schema



36

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
<http://tm.kit.edu/>

## Dynamische Kontrolle des Multicast-Baums (2)



### ■ Pruning

- Ziel: Multicast-Baum wird um nicht benötigte Zweige zurückgeschnitten
- Vorgehen
  - Router verwaltet Liste mit abhängigen Routern und Schnittstellen
  - Sind keine Gruppenmitglieder und keine abhängigen Router an einer Schnittstelle mehr vorhanden, kann er dem Router stromaufwärts eine Pruning-Nachricht schicken
  - Empfängt ein Router eine Prune-Nachricht, trägt er den stromabwärts gelegenen Router aus der Liste der abhängigen Router aus
  - Sind alle Einträge auf diese Weise aus der Liste entfernt worden, schickt der Router ebenfalls eine Prune-Nachricht an den Router stromaufwärts
  - Pruning gilt nur für eine bestimmte Zeit, danach wird wieder geflutet

37

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
http://tm.kit.edu/

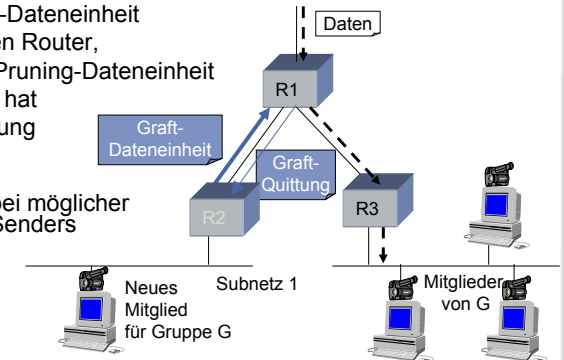
## Dynamische Kontrolle des Multicast-Baums (3)



### ■ Grafting

- Ziel: Wiedereingliederung eines Routers in den Multicast-Baum
- Ablauf
  - Tatsache, dass ein neues Gruppenmitglied in einem abgekoppelten Teilnetz vorhanden ist, wird über IGMP/MLD bekannt
  - Router sendet Graft-Dateneinheit an vorangegangenen Router, falls er vorher eine Pruning-Dateneinheit an diesen gesendet hat
  - Router sendet Quittung
    - Erkennen von Datenverlusten
    - Rückmeldung bei möglicher Inaktivität des Senders

### ■ Schema



38

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
http://tm.kit.edu/

### 6.6.3 MOSPF



- Multicast-Erweiterung von OSPF (Open Shortest Path First)
  - empfängerorientiertes Verfahren
  - basiert wie auch OSPF auf dem Link-State-Algorithmus
- Multicast-Erweiterungen von OSPF sind rückwärtskompatibel
  - MOSPF-Router sind interoperabel mit OSPF-Routern (Unicast-Verkehr)
    - RIP und DVMRP sind dagegen nicht interoperabel
  - Beim gemeinsamen Betrieb von OSPF- und MOSPF-Routern in einem Netz muss ein MOSPF-Router zum Designated-Router gewählt werden, damit Multicast-Verkehr in das entsprechende Subnetz weitergeleitet wird
- Erweiterungen
  - lokale Gruppenzugehörigkeit muss in den Routern bekannt sein
  - für jedes Paar aus Sender S und Gruppe G muss ein eigener Multicast-Baum berechnet werden

39

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
<http://tm.kit.edu/>

- MOSPF ist ein Link-State-Routing-Protokoll
- Router kennen die vollständige Netztopologie ihres Bereichs
- Aus Gruppendatenbasis und Netztopologie können die Router für jedes Sender-Gruppe-Paar Multicast-Bäume berechnen
- Kalkulation eines Multicast-Baumes wird nur bei Bedarf durchgeführt
- Berechnete Routing-Informationen werden in Forwarding-Caches zwischengespeichert
  - Eintrag umfasst
    - Schnittstelle zum Sender (Upstream)
    - Schnittstellen zu Empfängern (Downstream)
    - noch anfallende Kosten bis zu den Empfängern
- Bei Änderungen der Netztopologie: Forwarding-Caches löschen, Multicast-Bäume neu berechnen
- Multicast zwischen autonomen Systemen
  - Ähneln Multicast zwischen Bereichen
  - Aus der Menge der Grenz-AS-Router werden Inter-AS-Multicast-Forwarder bestimmt, die als Wildcard-Multicast-Empfänger operieren
  - Spezielles Protokoll für Multicast zwischen autonomen Systemen nötig
    - nicht Bestandteil von MOSPF
- Bewertung MOSPF
  - + Berechnung effizienter Verteilbäume
  - Schlechte Skalierbarkeit, da empfängerorientierte Zustandshaltung auch notwendig, wenn keine Sender vorhanden

### 6.6.4 PIM: Protocol Independent Multicast



- PIM unterstützt verschiedene Einsatzszenarien
  - Räumlich verstreute Gruppen mit geringer Dichte an Mitgliedern (**Sparse**)
  - Räumlich wenig verstreute Gruppe mit hoher Dichte an Mitgliedern (**Dense**)
- PIM
  - verwendet existierendes Unicast-Routingprotokoll zur Topologieerkennung
  - ist unabhängig vom konkreten Typ des Unicast-Routingprotokolls (deshalb „**Protocol Independent**“)
- Ziele
  - Minimierung der Zustandshaltung in den Routern
  - Minimierung des Verarbeitungsaufwands von Kontroll- und Nutzdaten
  - Minimierung der beanspruchten Bandbreite im Netz

40

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
<http://tm.kit.edu/>

- PIM verwendet existierendes Unicast-Routingprotokoll zur Topologieerkennung im Gegensatz zu DVMRP, das ein eigenes Protokoll dafür implementiert

## PIM-Varianten



### ■ PIM steht als Oberbegriff für unterschiedliche Protokolle

- **PIM-Dense Mode (PIM-DM)** [RFC3973]
  - Für Gruppen mit hoher Dichte
  - Basiert auf Fluten und Pruning
- **PIM-Sparse Mode (PIM-SM)** [RFC4601]
  - Für Gruppen mit niedriger Dichte
  - Basiert auf Rendezvous-Stellen (bildet Wurzel des Multicastbaums)
- **Bidirectional PIM (BIDIR-PIM)** [RFC5015]
  - Für Multipeer-Kommunikation am Besten geeignet
  - Bidirektionaler Baum mit Rendezvous-Stellen-Link als Wurzel des Baums
  - Sender befinden sich ebenfalls in diesem Baum: nur ein Baum für alle Sender einer Gruppe

41

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



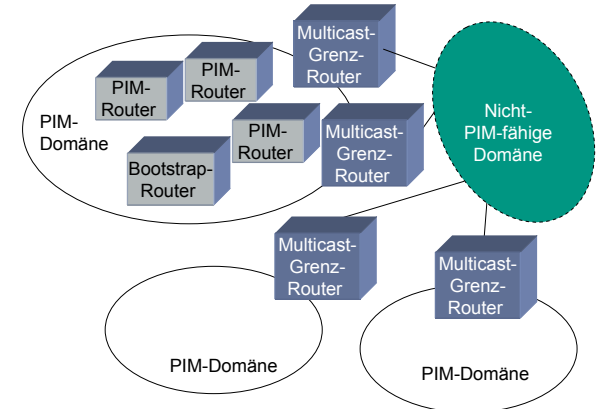
Institut für Telematik, Fakultät für Informatik  
<http://tm.kit.edu/>

## PIM-Architektur



### ■ Aufbau

- PIM-fähige Domänen
  - Mit **Bootstrap-Router**, der Information über Rendezvous-Stellen verbreitet
- Nicht PIM-fähige Domänen
- Verbindung von Domänen über Multicast-Grenz-Router



42

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
<http://tm.kit.edu/>

## PIM Sparse-Mode

### ■ Annahmen

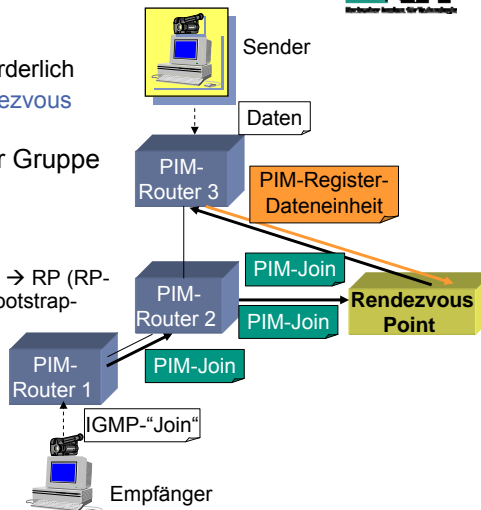
- Expliziter Gruppenbeitritt ist erforderlich
- Rendezvous-Stellen (oder **Rendezvous Point, RP**) werden bereitgestellt

### ■ Beitritt eines Empfängers zu einer Gruppe

- IGMP-Join im lokalen Subnetz
- Designated-Router (DR) sendet **PIM-Join** zur Rendezvous-Stelle
  - Abbildung MC-Gruppenadresse → RP (RP-set) vorkonfiguriert oder über Bootstrap-Nachrichten bekannt
  - PIM-Join installiert Zustand (\*,G) in durchquerten Routern
  - periodisches Senden von PIM-Join

### ■ Sender wird aktiv

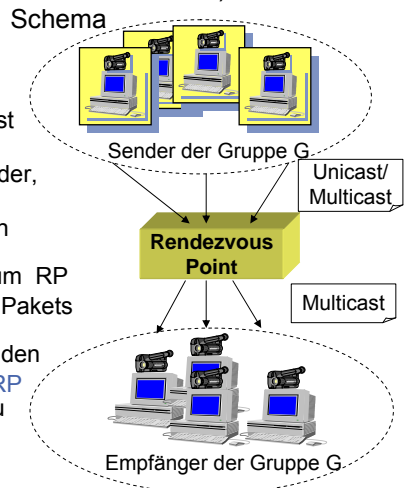
- Sender-DR kapselt Datenpaket in **PIM-Register-Paket**, wird per Unicast an Rendezvous-Stelle geschickt (**PIM-Register-Tunnel**)



## Prinzipieller Aufbau

### ■ Rendezvous-Stellen-Baum (RPT, Rendezvous Point Tree)

- Rendezvous-Stelle repräsentiert zentralen Knoten für die Empfänger
- RP ist Wurzel des gemeinsamen Baums, der von allen Quellen einer Gruppe genutzt wird, zuerst per Unicast im **PIM Register Tunnel**
- RP schickt **(S,G)-Join** in Richtung Sender, um Daten ohne Tunnel zu bekommen
- Sender-DR startet nach Erhalt des Join **natives Multicast** vom Sender zum RP entlang Sender-spezifischem Baum zum RP
- Nach Erhalt des ersten ungetunnelten Pakets (natives Multicast), schickt der RP ein **Register-Stop**, um Einkapseln zu beenden
- **Shortest-Path-Tree von Quellen zum RP** kann Multicast-Verteilbaum vom RP zu den Empfängern schneiden: Daten fließen ohne Umweg über RP in Verteilbaum, Sender wird aus Verteilbaum genommen ((S,G,RPT)-Zustand)



## Sender-spezifischer Baum (1)



### ■ Spezifischer Baum für Sender

- Source/sender-specific **Shortest Path Tree (SPT)**
- Behebt Ineffizienz bei ungünstiger Lage der Rendezvous-Stelle
- Entlastet Rendezvous-Stelle
- Übergang vom gemeinsam genutzten Baum zum SPT wird durch DR des Empfängers initiiert
  - z.B. bei Überschreiten einer bestimmten gemittelten Daten- bzw. Paketrage
  - Kann an der Rendezvous-Stelle vorbei führen

### ■ Nachteile

- Zusätzlicher Zustand (einen je Sender, anstatt einen je Gruppe)
- Kein Abbau des sender-spezifischen Baums definiert (prinzipiell aber möglich)

45

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



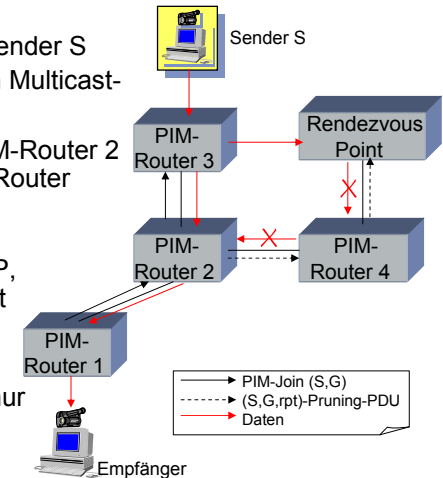
Institut für Telematik, Fakultät für Informatik  
http://tm.kit.edu/

## Sender-spezifischer Baum (2)



### ■ Beispiel

- PIM-Router 1 (DR) initiiert Übergang zum spezifischen Baum
- PIM-Router 1 sendet periodisch Join-Dateneinheit in Richtung Sender S
- PIM-Router 1, 2 und 3 erzeugen Multicast-Routingeintrag (S,G)
- Ist der Baum etabliert, erhält PIM-Router 2 Daten jetzt von einem anderen Router (PIM-Router 3) stromaufwärts
- PIM-Router 2 sendet daraufhin **(S,G,rpt)-Pruning** in Richtung RP, um die Daten nicht mehr doppelt zu erhalten



### ■ Soft-States

- Multicast-Routingeinträge sind nur für bestimmten Zeitraum gültig
- periodisches PIM-Join

46

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
http://tm.kit.edu/

## Gemeinsamer Baum mit Rendezvous-Stelle



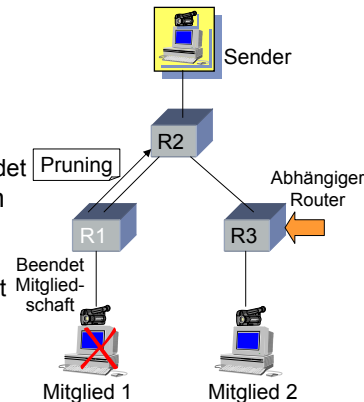
- In Routern zwischen Rendezvous-Stelle und Empfänger müssen Multicast-Routingeinträge vorhanden sein, um Datenpakete weiterzuleiten

### Zu berücksichtigende Situationen

- Eintrag zur Gruppe G mit beliebigem Sender vorhanden
  - Eintrag hat die Form (\*,G)
  - Gemeinsamer Baum wird verwendet
- Eintrag zur Gruppe G mit spezifischem Sender vorhanden
  - Eintrag hat die Form (S, G)
  - Spezifischer Baum wird verwendet

### Löschen eines Multicast-Routing-Eintrags

- keine lokalen Gruppenmitglieder mehr
- keine abhängigen Router



47

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
http://tm.kit.edu/

## Bootstrap-Router



### Aufgabe

- Verbreitung von Informationen über Rendezvous-Stellen innerhalb einer Domäne (Abbildung Gruppenadresse → RP, als **RP-Set** bezeichnet)
- Bootstrap Router (BSR)** Mechanismus für PIM [RFC5059]

### Ablauf

- Innerhalb einer Domäne werden einige Router als **Kandidaten für Bootstrap-Router (C-BSR)** und als **Kandidaten für Rendezvous-Stellen (C-RP)** konfiguriert
- Ein Kandidat wird als Bootstrap-Router ausgewählt
- Kandidaten für Rendezvous-Stellen melden sich bei Bootstrap-Router periodisch per Unicast an
  - Aus C-RP-Set wählt BSR die aktiven RPs aus (**RP-Set**)
- Bootstrap-Router versendet Bootstrap-Dateneinheiten mit RP-Set
  - Bekanntgabe der aktiven Rendezvousstellen innerhalb einer Domäne
- Router werten Bootstrap-Dateneinheiten aus, um eine passende Rendezvous-Stelle für eine Gruppe zu bestimmen

48

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
http://tm.kit.edu/



## BIDIR-PIM [RFC5015]



- Variante von PIM-SM, die bidirektionale Verteilbäume unterstützt
- PIM-SM erzeugt gemeinsam genutzte unidirektionale Verteilbäume
- Verteilen von Daten über gemeinsamen Verteilbaum mit zwei Modi:
  - Einkapseln durch Sender DR, übertragen an RP, dort entkapseln
  - Natives Multicast von Sender zum RP, erfordert sender-spezifischen Zustand in Routern
- BIDIR-PIM bedeutet Vereinfachung, da Sender-spezifische Zustände nicht mehr erforderlich sind
- Designated Forwarder (DF) entlang eines jeden Links des Verteilbaums
- Wahl des DF erfolgt automatisch

49

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
<http://tm.kit.edu/>

## PIM-Dense Mode [RFC3973]



- Nach langer Zeit als RFC 3973 (Experimental) neu verfasst, ursprüngliches Konzept bereits in S. Deerings Doktorarbeit von 1991
- Annahmen
  - Gruppenmitglieder befinden sich in fast allen Subnetzen („**dichte Empfänger**menge“)
- Mechanismus
  - **Fluten** (mit Reverse-Path-Forwarding Überprüfung) und **Pruning** (ähnlich wie DVMRP)
  - Keine Kenntnis über abhängige Router
  - Neue Zweige können mit Graft Messages installiert werden
  - Pruning-Zustand ist mit (S, G) assoziiert
  - Verwendung gleicher Paket-Formate wie bei PIM-SM
- Nachbarschaft
  - Periodische Hello-Nachrichten (wie bei PIM-SM, ident. Nachrichtenformat)
- Unterschied zu DVMRP
  - Unabhängig von Prozeduren, die Topologie herausfinden
  - Benutzung eines existierenden, beliebigen Unicast-Routingprotokolls (deswegen Protocol-Independent)

50

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
<http://tm.kit.edu/>

## 6.6.5 Multicast Inter-Domain Routing



- Eignung der zuvor vorgestellten Protokolle für Inter-Domain Multicast?
  - DVMRP und PIM-DM fluten periodisch (schlecht skalierbar)
  - MOSPF benötigt Kenntnis über die Lokation der Empfänger (schlecht skalierbar)
  - PIM-SM
    - erfordert Kenntnis über die Menge der Rendezvous-Stellen
    - erfordert global gesehen viele Rendezvous-Stellen zwecks Lastverteilung
    - Umgekehrter Unicast-Pfad für Multicast-Forwarding ist problematisch wenn nicht alle Router PIM-SM-fähig sind
    - Konnektivitätsverlust durch Ausfall einer Rendezvous-Stelle gefährdet Unabhängigkeit der ISPs



## Multicast Inter-Domain Routing – BGMP



- **BGMP (Border Gateway Multicast Protocol)**
  - Inter-Domain Multicast Protokoll mit gewissen Ähnlichkeiten zu BGP
  - Unterstützung
    - unidirektionale quellenbasierte Verteilbäume
    - unidirektionale geteilte Verteilbäume
    - bi-direktionale geteilte Verteilbäume
  - Geteilte Bäume haben Wurzel in Autonomem System (AS): erfordert Zuordnung von Multicast-Adressen zu ASen (z.B. Multicast-Adressenprotokoll MASC)
  - Zuordnung wird mittels MBGP (Multiprotocol Extensions for BGP-4) verteilt



## Multicast Inter-Domain-Routing – MSDP



### ■ Multicast Source Discovery Protocol (MSDP)

- Übergangsstrategie, keine langfristige Lösung
- Domänen verwenden PIM-SM, haben vollständige Menge (d.h. für alle Multicast-Gruppen) von Rendezvous-Stellen
- MSDP ermöglicht lockere Vermaschung der Rendezvous-Stellen zwischen den Domänen
- Wird ein Sender in einer Domäne aktiv, werden alle über MSDP verbundenen Rendezvous-Stellen mit einer „Source-Active“-Nachricht benachrichtigt
- Benachbarte Rendezvous-Stellen schicken sender-spezifisches Join in Richtung des Senders
- Nachteil: jede Rendezvous-Stelle muss über Senderaktivitäten benachrichtigt werden, daher schlecht skalierbar
- Nur für IPv4 spezifiziert!

53

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
http://tm.kit.edu/

## Embedded RP Address



- Problem: MSDP nicht für IPv6 definiert, wie können Sender zwischen Domänen (Inter-domain) gefunden werden?
- Einfache Lösung für IPv6: Einbetten der RP-Adresse in IPv6-Multicast-Adresse
- Embedded RP Address [RFC3956]

11111111	Flags	Scope	Re-served	RIID	Plen	Network Prefix	Group ID
8	4	4	4	4	8	64	32 Bits

0|R|P|T

R=1: Multicast-Adresse enthält eingebettete RP-Adresse  
RIID: RP Interface ID

- Aus Multicast-Adresse **abgeleitete RP-Unicast-Adresse:**

Network Prefix	RIID
Plen	4 Bits
00000000000000000000000000000000	

54

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
http://tm.kit.edu/

- RIID ist notwendig, damit man nicht die Anycast-Adresse des für das Netzpräfix zuständigen Routers erwischt.

## Anycast-RP – Motivation



### ■ Probleme mit PIM-SM:

- Nur ein aktiver RP pro Multicast-Gruppe (je Domäne) möglich, dadurch ggf. auch topologisch ineffiziente Szenarien vorhanden
- Rendezvous-Point stellt Single-Point-of-Failure dar, bei Ausfall langsame Konvergenz
- Lastverteilung (insbesondere für Entkapseln) zwischen Rendezvous-Points

### ■ Lösung

- mehrere RPs einsetzen
- Menge von RPs als **Anycast-RP-Set** [RFC4610]

55

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)

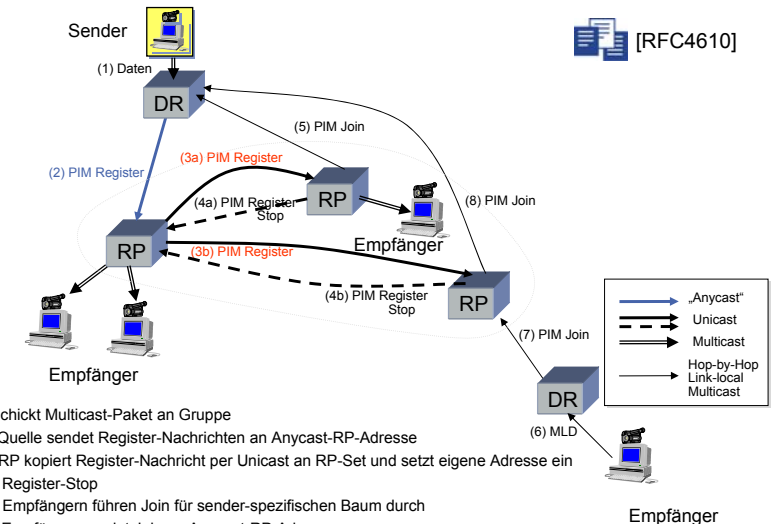


Institut für Telematik, Fakultät für Informatik  
http://tm.kit.edu/

## Anycast RP – Funktionsweise



[RFC4610]



- (1) Quelle schickt Multicast-Paket an Gruppe
- (2) DR der Quelle sendet Register-Nachrichten an Anycast-RP-Adresse
- (3) „erster“ RP kopiert Register-Nachricht per Unicast an RP-Set und setzt eigene Adresse ein
- (4) optional Register-Stop
- (5) RPs mit Empfängern führen Join für sender-spezifischen Baum durch
- (6-7) neuer Empfänger sendet Join an Anycast-RP-Adresse
- (8) RP kennt Quelle bereits und führt sender-spezifisches Join sofort durch

56

Next Generation Internet SS2010 – 6. Multicast-Routing (R0)



Institut für Telematik, Fakultät für Informatik  
http://tm.kit.edu/

- Ziel: keine Änderung der PIM-SM-Funktionsweise
- Nutzung von Anycast
  - Empfänger senden Join an Anycast-RP-Adresse (per Link-local Multicast Hop-by-Hop in Richtung RP, Route durch Anycast vorgegeben)
  - DRs der Quelle senden Register-Nachrichten an Anycast-RP-Adresse
  - Unicast-Routing leitet dann jeweils an nächstgelegenen RP weiter
- jeder RP kennt Unicast-Adressen aller anderen RPs des Anycast-RP-Sets
- RP kopiert ankommendes Register-Paket an alle bekannten RPs
  - trägt dabei seine Unicast-Adresse als Quelladresse ein
  - daher verteilen die übrigen RPs dieses Register nicht weiter an Anycast-RP-Set
  - dieser RP wird üblicherweise dann auch Join zu senderspezifischem Baum durchführen
- Die übrigen RPs erkennen, dass Sender aktiv wurde
  - etablieren Zustand für die Quelle
  - führen sender-spezifisches Join durch falls Empfänger vorhanden
- Für Anycast-RP dient Register lediglich zur Erkennung von Quellen, nicht mehr zur Transport der Daten
- Frage: woher ist Anycast-RP-Adresse für Multicast-Gruppe bekannt? (wie üblich: Embedded RP oder BSR)
- Anycast-Routing erfordert das Eintragen der Routen zu den RPs
- RP kennt Unicast-Adressen aller anderen seiner Domäne (z.B. manuell konfiguriert)
- Register-Stop zwischen RPs kann gesendet werden, ist aber wirkungslos (es werden Register-Nachrichten weiterhin geschickt)
- Wieso dient Register im Wesentlichen nur noch zum Erkennen der Quellen: Nur erster RP hat Einfluss darauf, wann Register-Stop gesendet wird, die anderen RPs wären dann von Datenfluss abgeschnitten. Deshalb sollten sie umgehend ein sender-spezifisches Join durchführen.
- Daten können während der Join-Phase kurzzeitig verloren gehen
- Vereinfachter Anycast-RP-Mechanismus aus RFC4610 benötigt auch kein MSDP mehr, im Gegensatz zu RFC 3446

## 6.7 Übungen



- 6.1 Welche Aufgabe haben Designated-Router?
- 6.2 Können mit IGMP bekannte Gruppen auf der Sicherungsschicht realisiert werden? Begründung.
- 6.3 Welchen Vorteil bringt die Pruning-Technik mit sich?
- 6.4 Welche Probleme bringen Rendezvous-Stellen mit sich?
- 6.5 Inwiefern ist PIM Protokoll-unabhängig?
- 6.6 Nennen Sie die Hauptunterschiede von PIM sparse mode und PIM dense mode.
- 6.7 Zu welchem Zweck können Anycast-Adressen eingesetzt werden?



## 6.9 Referenzen und weiterführende Literatur



- [BCHC02] I. Brown, J. Crowcroft, M. Handley, B. Cain, Internet Multicast Tomorrow, Internet Protocol Journal, Dezember 2002, [http://www.cisco.com/warp/public/759/ipj\\_5-4/ipj\\_5-4\\_internet\\_multicast.html](http://www.cisco.com/warp/public/759/ipj_5-4/ipj_5-4_internet_multicast.html)
- [RFC 1112] S.E. Deering. Host extensions for IP multicasting. RFC 1112 (Standard), August 1989. Updated by RFC 2236. URL: <http://www.ietf.org/rfc/rfc1112.txt>
- [RFC 2189] A. Ballardie. Core Based Trees (CBT version 2) Multicast Protocol Specification. RFC 2189 (Experimental), September 1997. URL: <http://www.ietf.org/rfc/rfc2189.txt>
- [RFC 2201] A. Ballardie. Core Based Trees (CBT) Multicast Routing Architecture. RFC 2201 (Experimental), September 1997. URL: <http://www.ietf.org/rfc/rfc2201.txt>
- [RFC 2236] W. Fenner. Internet Group Management Protocol, Version 2. RFC 2236 (Proposed Standard), November 1997. Obsoleted by RFC 3376. URL: <http://www.ietf.org/rfc/rfc2236.txt>
- [RFC 2365] D. Meyer. Administratively Scoped IP Multicast. RFC 2365 (Best Current Practice), Juli 1998. URL: <http://www.ietf.org/rfc/rfc2365.txt>



## 6.9 Referenzen (2)



- [RFC 2730] S. Hanna, B. Patel und M. Shah. Multicast Address Dynamic Client Allocation Protocol (MADCAP). RFC 2730 (Proposed Standard), Dezember 1999. URL: <http://www.ietf.org/rfc/rfc2730.txt>
- [RFC 2908] D. Thaler, M. Handley und D. Estrin. The Internet Multicast Address Allocation Architecture. RFC 2908 (Informational), September 2000. URL: <http://www.ietf.org/rfc/rfc2908.txt>
- [RFC 2909] P. Radoslavov, D. Estrin, R. Govindan, M. Handley, S. Kumar und D. Thaler. The Multicast Address-Set Claim (MASC) Protocol. RFC 2909 (Experimental), September 2000. URL: <http://www.ietf.org/rfc/rfc2909.txt>
- [RFC 3306] B. Haberman und D. Thaler. Unicast-Prefix-based IPv6 Multicast Addresses. RFC 3306 (Proposed Standard), August 2002. Updated by RFCs 3956, 4489. URL: <http://www.ietf.org/rfc/rfc3306.txt>
- [RFC 3376] B. Cain, S. Deering, I. Kouvelas, B. Fenner und A. Thyagarajan. Internet Group Management Protocol, Version 3. RFC 3376 (Proposed Standard), Oktober 2002. Updated by RFC 4604. URL: <http://www.ietf.org/rfc/rfc3376.txt>



## 6.9 Referenzen (3)



- [RFC 3446] D. Kim, D. Meyer, H. Kilmer und D. Farinacci. Anycast Rendezvous Point (RP) mechanism using Protocol Independent Multicast (PIM) and Multicast Source Discovery Protocol (MSDP). RFC 3446 (Informational), Januar 2003. URL: <http://www.ietf.org/rfc/rfc3446.txt>
- [RFC 3956] P. Savola und B. Haberman. Embedding the Rendezvous Point (RP) Address in an IPv6 Multicast Address. RFC 3956 (Proposed Standard), November 2004. URL: <http://www.ietf.org/rfc/rfc3956.txt>
- [RFC 3973] A. Adams, J. Nicholas und W. Siadak. Protocol Independent Multicast - Dense Mode (PIM-DM): Protocol Specification (Revised). RFC 3973 (Experimental), Januar 2005. URL: <http://www.ietf.org/rfc/rfc3973.txt>
- [RFC 4291] R. Hinden und S. Deering. IP Version 6 Addressing Architecture. RFC 4291 (Draft Standard), Februar 2006. URL: <http://www.ietf.org/rfc/rfc4291.txt>
- [RFC 4601] B. Fenner, M. Handley, H. Holbrook und I. Kouvelas. Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised). RFC 4601 (Proposed Standard), August 2006. URL: <http://www.ietf.org/rfc/rfc4601.txt>



## 6.9 Referenzen (4)



- [RFC 4610] D. Farinacci und Y. Cai. Anycast-RP Using Protocol Independent Multicast (PIM). RFC 4610 (Proposed Standard), August 2006. URL: <http://www.ietf.org/rfc/rfc4610.txt>
- [RFC 5015] M. Handley, I. Kouvelas, T. Speakman und L. Vicisano. Bidirectional Protocol Independent Multicast (BIDIR-PIM). RFC 5015 (Proposed Standard), Oktober 2007. URL: <http://www.ietf.org/rfc/rfc5015.txt>
- [RFC 5059] N. Bhaskar, A. Gall, J. Lingard und S. Venaas. Bootstrap Router (BSR) Mechanism for Protocol Independent Multicast (PIM). RFC 5059 (Proposed Standard), Januar 2008. URL: <http://www.ietf.org/rfc/rfc5059.txt>
- [WiZi99] R. Wittmann, M. Zitterbart, Multicast: Protokolle und Anwendungen, dpunkt.Verlag, 1999, ISBN: 3-920993-40-3
- [WiZi00] R. Wittmann, M. Zitterbart, Multicast Communication: Protocols and Applications, Morgan Kaufman Publisher, 2000, ISBN: 1-55860-645-9

