

Internet – Protokolle für Multimedia - Anwendungen

Kapitel 5.2 IP Multicast Routing

Gliederung

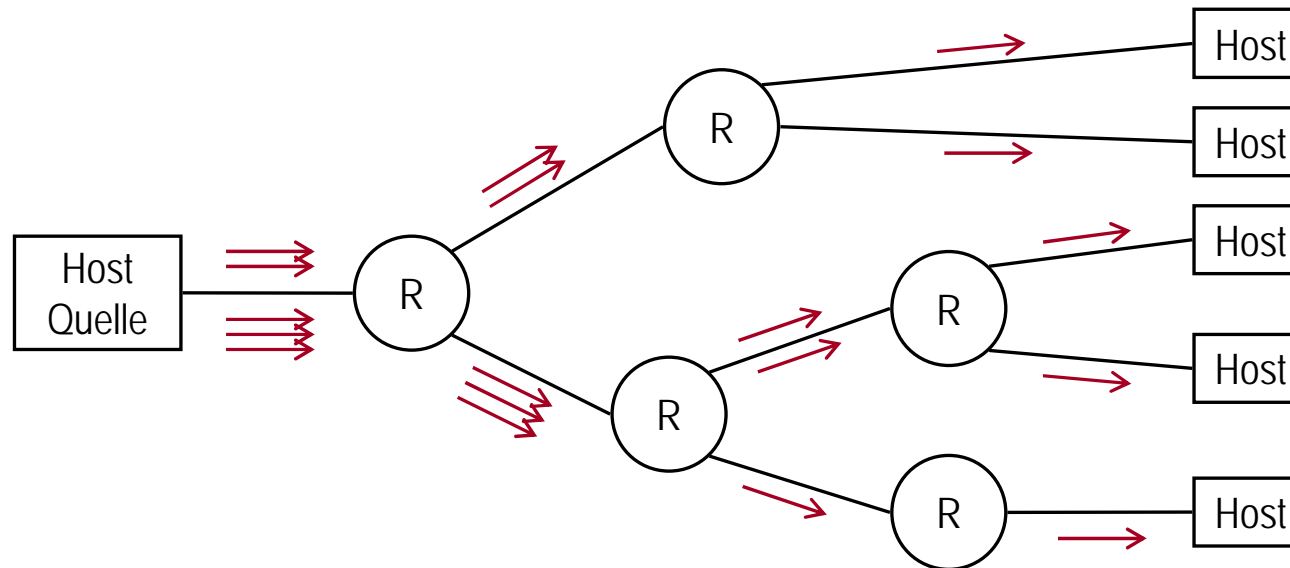
- Warum Multicast?
- IP-Multicasting
- Internet Groups Management Protocol IGMP
- Multicast Routing Protokolle
 - MBONE
 - Distance Vector Multicast Routing Protocol DVMRP
 - Multicast OSPF
 - Protocol Independent Multicast, PIM-SM
 - Core Based Tree

Warum Multicast? (1)

- ❑ Viele MM-Anwendungen (z.B. Video-Verteilung, Videokonferenzen) sind Gruppenkommunikationen, also n:m-Beziehungen.
- ❑ Emulation durch Unicast, also 1:1-Beziehungen, prinzipiell möglich, erzeugt aber viele Kopien, u.U. auf gleichen Strecken
- ❑ Multicast gestattet die Kopierzeugung erst im letzten Moment, außerdem skaliert Unicast nicht
- ❑ Multicast somit sinnvoll
 - für bessere Bandbreitenausnutzung
 - Weniger Router-/Host-Processing
 - falls Empfängeradressen nicht bekannt

Warum Multicast? (2)

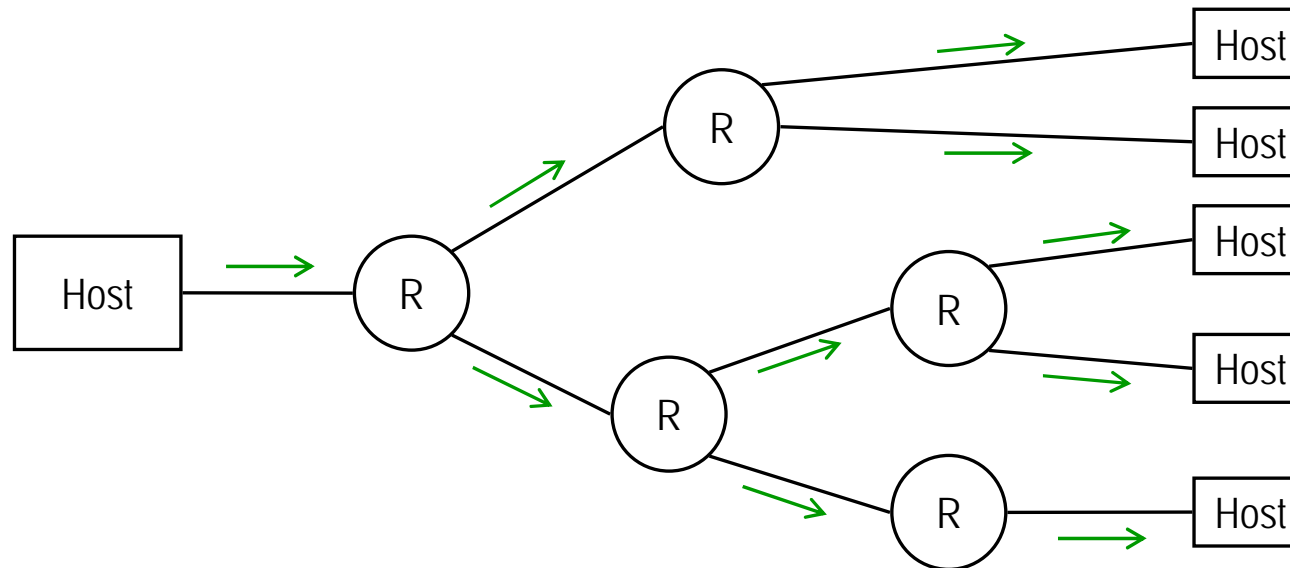
□ Unicast



Für jede Quell-Senkenbeziehung wird ein vollständiger Datenfluss benötigt.

Warum Multicast? (3)

□ Multicast



Datenkopien werden erst dann erzeugt, wenn sie an Verzweigungen erforderlich werden.

IP – Multicasting (1)

- ❑ Verwendet wird die Internet-Adressklasse D



umfasst Adressraum 224.0.0.0 bis 239.255.255.255, somit ca. 250 Millionen Adressen

- ❑ Empfänger einer Sendung soll eine Multicast-Gruppe sein. Quelle selbst muss nicht Mitglied sein
- ❑ Problemfelder sind
 - Adressen-Mapping: Abbilden von IP-Multicast auf MAC-Multicast
 - Endsystem-Registrierung: Empfänger muss Gruppe angehören. Registrierung mittels IGMP
 - Multicast Routing: Informationen über Gruppen müssen verteilt und gepflegt werden. Dazu Protokolle PIM, MOSPF, DVMRP, MBONE

IP – Multicasting (2)

- ❑ Dedizierte Multicast-Adressen in der D Class (RFC 1700/IANA) reserviert sind
 - 224.0.0.0 bis 224.0.0.255 als Local Scope Addresses, die nicht global geroutet werden.
 - All-Host-Group 224.0.0.1
 - All-Router-Group 224.0.0.2 mit Unterteilung
 - all DMVRP-Router 224.0.0.4
 - all OSPF-Router 224.0.0.5
 - all OSPF-Designated Routers 224.0.0.6
 - all PIM-Routers 224.0.0.13
 - Applikationsspezifische Multicasts (Auswahl)
 - Network Time Protocol 224.0.1.1
 - IETF-1-Audio 224.0.1.11
 - IETF-2-Video 224.0.1.15

IP – Multicasting (3)

□ Dedizierte Multicast-Adressen

- Global Scope Addresses
 - Adressen 224.0.1.0 bis 238.255.255.255
 - werden nur temporär zugewiesen
- Administratively Scoped Addresses
 - Adressen 239.0.0.0 bis 239.255.255.255
 - davon Site-local scope: 239.253.0.0/16
 - Organization-local scope: 239.192.0.0/14

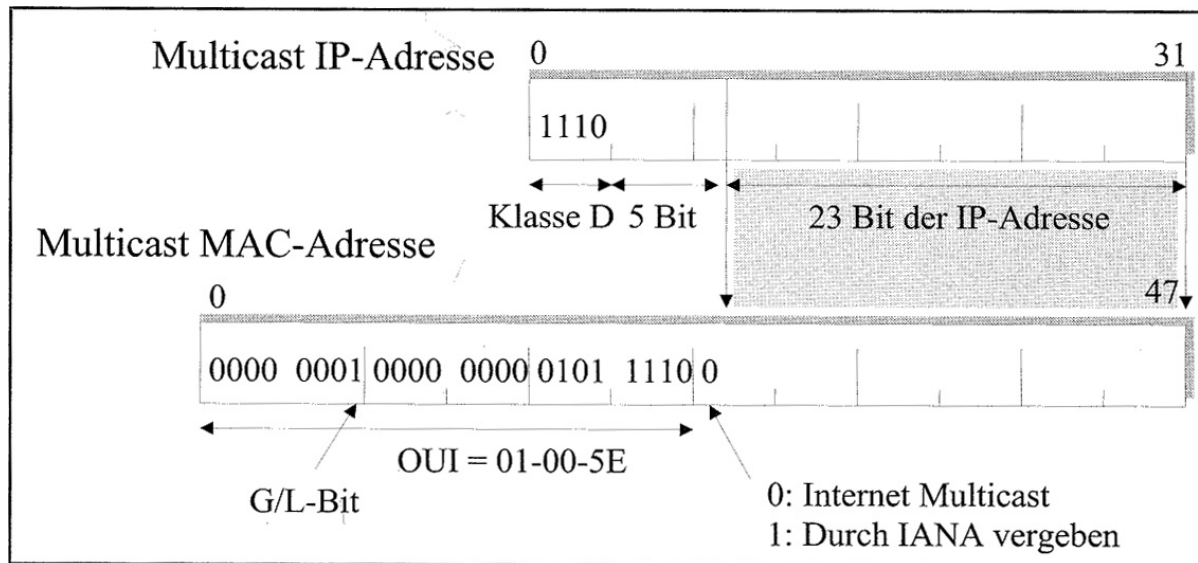
IP-Multicasting (4)

- Nicht-reservierte Multicast-Adressen werden temporär für Gruppen eingesetzt. Dadurch werden einem IP-Interface zusätzlich zu seiner Unicast-Adresse u.U. mehrere Multicast-Adressen zugeordnet
- Eine MC-Adresse steht für Anzahl von k Endsystemen. Alle diese Systeme gehören zur All-Host-Group, somit kann der zugehörige M-Router über 244.0.0.2 erreicht werden
- M-Router verwaltet nicht eine zusätzliche $k:1$ Adresstabelle für die MAC-Adressen, sondern MC-fähige Hosts verwalten neben MAC-Unicast-Adresse auch MAC-MC-Adresse
- Abbildung MC-IP auf MC-MAC über OUI und G/L-Bit. G/L (Group Local) kennzeichnet MC-MAC. 23 Bit bleiben für Adress-Abbildung, dadurch mögliche Überschneidungen
- Über TTL-Bit kann MC-Reichweite gesteuert werden.

IP – Multicasting (5)

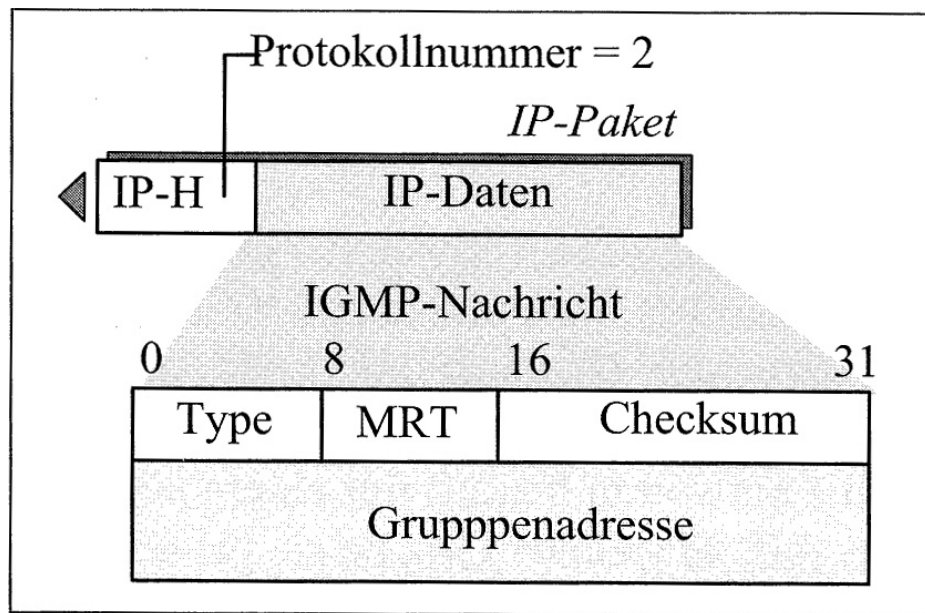
□ Mapping von IP-auf MAC-Multicast-Adresse

OUI = Organisation Unique Identifier, G/L = Group/Local Bit



IGMP (1)

- ❑ IGMP (Internet Group Management Protocol) RFC 2236: IGMPv2
- ❑ baut auf IP auf und dient der Multicast-Gruppenverwaltung



IGMP (2)

□ IGMP-Nachrichten

- Type (für Host-to-M-Router Kommunikation)
 - Membership Query (x'11'): Welche Gruppen gibt es? Welche Mitglieder gibt es in bestimmter Gruppe?
 - DVMRP-V3-Nachricht (0x'13')
 - Membership-Report (x'16'): zur Aufnahme von Endsystemen in Multicast-Gruppen (Membership Request)
 - Leave Group (x'17')
- Maximum Response Time (MRT): Hinweis zum Router-Tabellen-Update
- Checksum: über gesamte IGMP-Nachricht
- Gruppenadresse: (in Anfrage = 0)

IGMP (3)

□ IGMP-Arbeitsweise

- M-Router sendet periodisch (100s) IGMP-PDUs vom Typ Query an alle Hosts im LAN über 224.0.0.1 mit TTL = 1. Alle Hosts antworten mit Membership Report und teilen mit, in welcher Gruppe sie sind.
- Soll Host eingefügt werden, sendet er MAC-Multicast mit IGMP Host Membership Report. Zuordnung MAC-IP-Adressen wie oben
- Will Host die Gruppe verlassen, analog mit Leave Group PDU. Anschließend antwortet Host nicht mehr auf Query
- M-Router muss Buchführung machen und Änderungen anderen Routern über Routingprotokolle mitteilen
- Da nicht alle Router M-Router sind, werden zwischen M-Routern oft Tunnel aufgesetzt.

Multicast – Routingprotokolle

- ❑ Ziel ist, Multicast-Verteilbäume auf der irregulär vermaschten Internettopologie aufzusetzen
 - Shortest Path oder Source Distribution Trees
 - Shared Distribution Trees
- ❑ Varianten, solche Bäume aufzubauen zu Multicast-Gruppen
 - MBONE: nutzt IP-Tunneling und DVRMP/RIP
 - DVRMP: nutzt RIP-ähnliche Multicast Routingtabellen und Poison-Reverse-Metrik, unterstützt auch Inter-Area-Routing
 - PIM: nutzt existierende Unicast Routing Tabellen und zusätzliche Join/Prune/Graft-Nachrichten. Intra-Area
 - MOSPF: nutzt unterliegendes OSPF. Intra-Area
 - CBT: ähnlich PIM

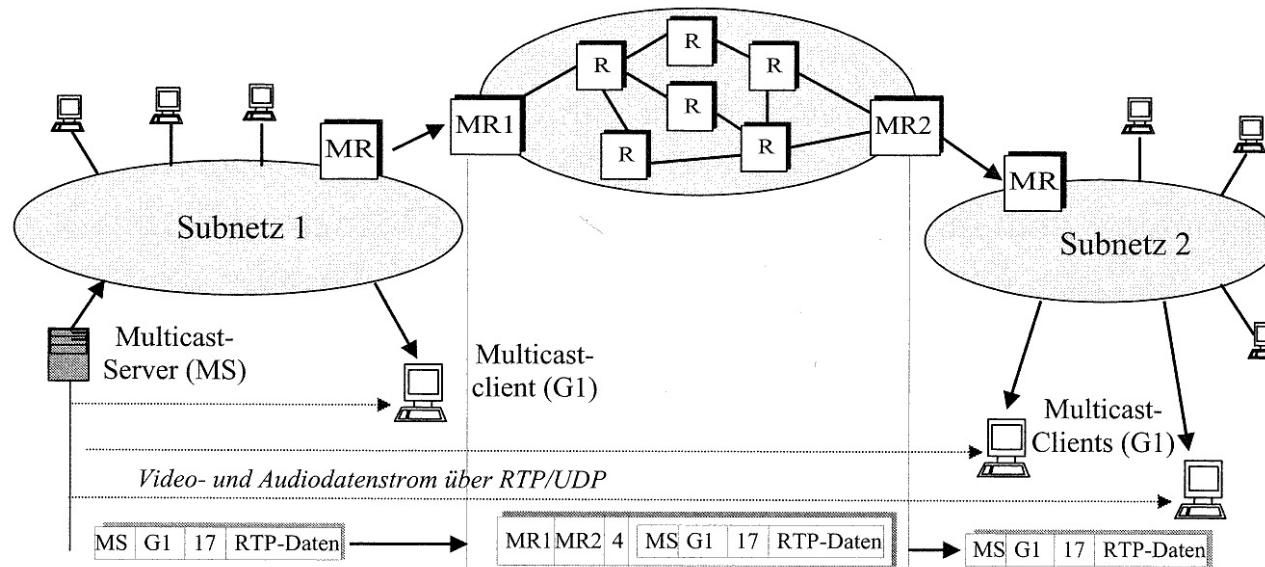
MBONE (1)

- ❑ MBONE (Multicast Backbone on the Internet). Frühe Unterstützung von A/V-Konferenzen. Definiert ein eigenes virtuelles Netz auf dem Internet. Realisiert in Bezug auf Multicast ein Autonomes System mit M-Routern. Besonders von Universitäten genutzt (über 5000 Subnetze)
- ❑ M-Router-Funktionen
 - Tunneling: IP-Paket mit UDP/RTP-Nutzlast bekommt beim Versand über Unicast-Router eine IP-IP-Encapsulation
 - M-Router verwenden DVMRP zwischen M-Router und RIP zwischen M-Routern und Unicast Routern
 - Reichweitenbegrenzung über TTL (TTL-Scoping), z.B. 0 (Knoten), 1 (Subnetz), 32 (Domain, e.g. DE), 48 (Europa), 64 (Region), 128 (Kontinent), 255 (unbegrenzt)

MBONE (2)

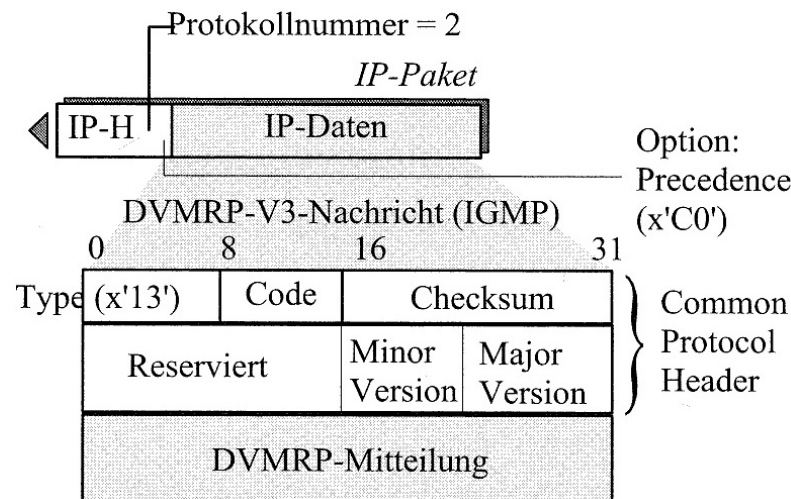
□ Arbeitsweise mit IP-IP-Encapsulation

R: Unicast Router, MRx: Multicast Router, G1: Multicast Group 1



DVMRP (1)

- ❑ DVMRP (Distance Vector Multicast Routing Protocol) benutzt IGMP. Gut für „dense“, d.h. dichtes Zusammenliegen der zu versorgenden Hosts (ineffektiv für „sparse“).
- ❑ DVMRPv3-Nachrichten.



DVMRP (2)

□ PDU-Aufbau (Version 3)

- Code: zeigt DVMRP-Kommandos an

Code	Kommando/ Nachrichtentyp	Beschreibung
1	Probe	Nachbarschafts-Entdeckung
2	Report	Austausch der Routing-Tabellen
5	Ask Neighbors2	Anfrage der Router-Nachbarschaftsliste (Unicast)
6	Nighbors2	Nachbar-Router und deren Eigenschaften (Unicast)
7	Prune	Pruning von Multicast-Pfaden
8	Graft	Grafting von Multicast Pfaden
9	Graft Ack	Acknowledgement für Graft-Nachrichten

DVMRP (3)

- DMVRP-Nachrichten a) Neighbor Probe, b) Reports für Routing
 T: Type, C: Code, SM: Subnet Mask
 (zur Abfrage von Router Capabilities und Status)

0	8	16	31
T=x'13'	C=x'1'	Checksum	
Reserved	Capabilities	Minor	Major
Generation ID			
Neighbor IP Address 1			
.....			
Neighbor IP Address n			

a)

0	8	16	31
T=x'13'	C=x'2'	Checksum	
Reserved		Minor	Major
SM1/3	SM2/3	SM3/3	SN1/3
SN2/3	SN3/3	Metric

b)

DVMRP (4)

- DMVRP-Nachrichten a) Prune, B) Graft
(zur dynamischen Verwaltung des Multicast Tree)

0	8	16	31
T=x'13'	C=x'7'	Checksum	
Reserved		Minor	Major
Source Host Address			
Group Address			
Prune Lifetime			
Source Network Mask			

a)

0	8	16	31
T=x'13'	C=x'7'	Checksum	
Reserved		Minor	Major
Source Host Address			
Group Address			
Source Network Mask			

b)

DVMRP – V3 – Routing (1)

- ❑ realisiert einen Reverse-Path-Multicast Algorithmus (Flood and Prune)
- ❑ Grundidee
 - (1) Sender schickt Multicast-Nachricht per Flooding
 - (2) Intermediate Router empfängt Nachricht mehrmals über verschiedene Wege
 - alle Pakete, die nicht über das Interface kommen, über das der Router Unicast an den Sender schicken würde, werden verworfen
 - Jede MC-PDU wieder an alle Interfaces außer an Quell-Interface
 - Somit unterliegendes Unicast Routing Protokoll erforderlich
 - (3) Router, die keine Multicast-Empfänger kennen, d.h. keine IGMP-PDUs empfangen haben, senden Prune-PDUs, um unnötige Nachrichten zu stoppen
 - (4) Später können neue Mitglieder über Graft-PDUs eingebunden werden.

DVMRP – V3 – Routing (2)

□ Routing-Details

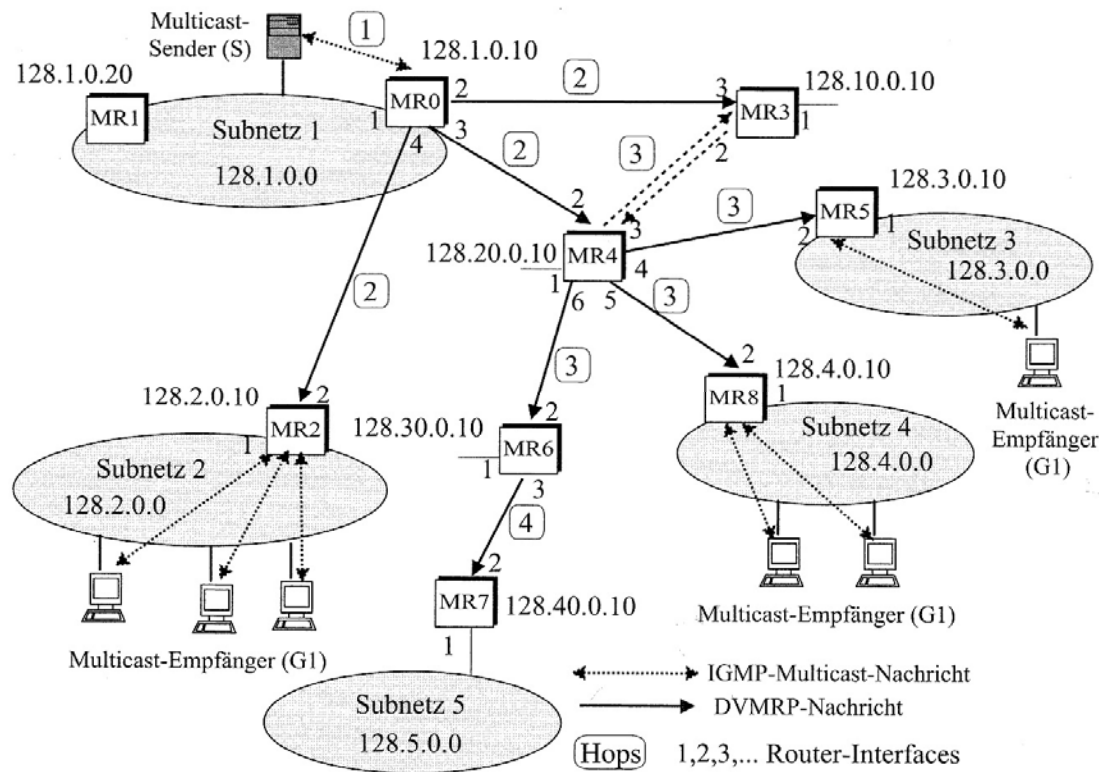
- Jedem Multicast Server, identifiziert durch seine MC-Adresse, wird ein Designated M-Router (DMR) zugewiesen, der Wurzel des M-Tree bildet. Jeder M-Gruppe ist ein eigener M-Tree zugewiesen.
- DMR hat pro Interface Verteilerlisten mit Mitgliedern definierter M-Gruppen
- Startannahme ist „dense mode“, also jeder Host im eigenen Netz ist Mitglied einer M-Group. Dann Anwenden M-Query über IGMP
- Ansprechen von M-Router über periodische Neighbor Probes per All-Router-MCs mit TTL=1 und Eintrag in lokale Verteilerlisten. Stufenweiser Aufbau des Baumes. MCs werden nur akzeptiert, wenn Interface auch zum Unicast passt.
- Im Gegensatz zu RIP wird hier Reverse Path Forwarding (RPF) genutzt. Aufbau vom Empfänger zum Multicast-Sender

DVMRP – V3 – Routing (3)

- ❑ Leaf Router: Router, die schließlich die MC-PDU in ein Subnetz ausliefern (z.B. über MAC-MC)
- ❑ Infinity Metrik: Maß für Baumlänge. Größerer Wert als der aktuelle bedeutet neu hinzukommendes Subnetz. Ansonsten Entfernung (Poison Reverse) mittels Truncated Reverse Path Broadcast
- ❑ Pruning: Abhängen eines Teil-Pfades aus M-Tree, falls auf Interface keine M-Gruppe mehr zu unterstützen
- ❑ Grafting: Wiederherstellen eines entfernten Pfades durch M-Router. Nach Empfang einer IGMP-Member-Query
 - Prüfen, ob er vom Tree entfernt wurde (Prune)
 - Prüfen, ob der Upstream-Router Prune/Graft-PDUs verarbeiten kann
 - Senden Graft-PDU zum Upstream-Router

DVMRP - V3 - Routing (4)

Beispiel eines Multicast-Routing-Baumes (1)



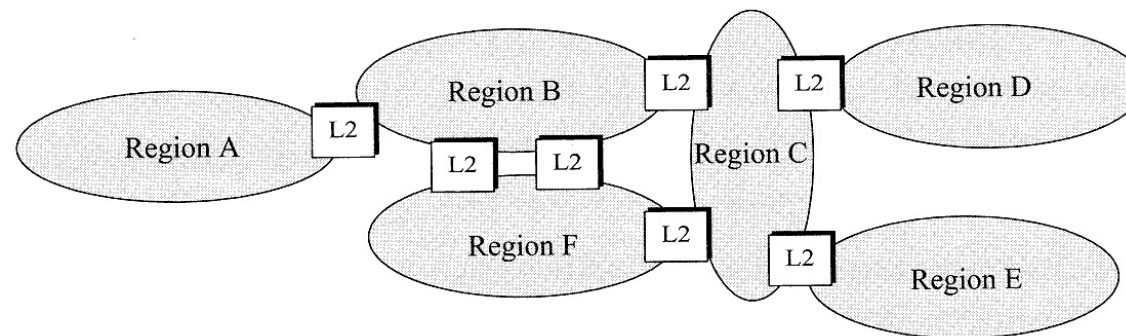
DVMRP – V3 – Routing (5)

Beispiel eines Multicast-Routing-Baumes (2): Routing-Tabellen

M-Router	Quellnetz	Von M-Router	Metrik	TTL	InPort	OutPort
MR0	128.1.0.0		1	200	1	0
MR1	128.1.0.0		1	150	1	2,3,4
MR2	128.1.0.0	128.1.0.10	2	150	2	1
MR3	128.1.0.0	128.1.0.10	2	150	3	1,2
MR4	128.1.0.0	128.1.0.10	2	100	2	1,3,4,5,6
MR5	128.1.0.0	128.20.0.10	2	50	2	1
MR6	128.1.0.0	128.20.0.10	3	50	2	1,3
MR7	128.1.0.0	128.30.0.10	4	50	2	1

DVMRP - V3 - Routing (6)

- ❑ Hierarchical DVMRP als Erweiterung zum Interdomain Routing
- ❑ Unterscheidung Level 1- und Level 2- M-Router. Jede Domain wird mit einem Region Identifier benannt, der auf Level 2 mitgeteilt wird.
- ❑ Intradomain wird DVMRP (oder auch MOSPF) eingesetzt, Interdomain HDVMPRP. L2-Router müssen beides können.



Multicast OSPF

- ❑ Erweiterung von OSPF zu MOSPF durch zusätzliches Multicast Link State Advertisement (LSA)
- ❑ Die OSPF-Datenbank beschreibt in jedem M-Router mit MOSPF die gesamte Topologie des Autonomen Systems einschl. Pfadkosten. Konsistenz wird über MC-Gruppen LSA und IGMP erreicht (Group Membership)
- ❑ Multicast Tree wird als Minimal Spanning Tree für Adresspaar (MC-Sender, MC-Gruppe) gebildet: per source / per group multicast tree
- ❑ Koexistenz von OSPF-Router und MOSPF-Router, für MC nur letztere zuständig
- ❑ Auch Inter-Area Multicasting möglich über Wild-Card M-Receiver und einem ausgezeichneten MOSPF-Router pro Area

Protocol Independent Multicast PIM (1)

- ❑ Annahme für DVMRP und MOSPF: dichte (dense) M-Gruppen
 - ❑ Bei dense-mode: Wenige M-Router halten Tabellen für beieinander liegende Gruppen
bei sparse-mode: Viele M-Router halten Tabellen für viele Gruppen mit wenigen lokalen Empfängern
- PIM unterstützt beide Modi: PIM-DM und PIM-SM
- ❑ PIM arbeitet unabhängig von dem zugrundeliegenden Unicast-Routing und hält eigene Multicast Routing Information Base (MRIB)
 - ❑ PIM setzt voraus das zwischen Subnetzen MC-IP-Pakete gesendet werden können, wobei sich M-Sender und M-Empfänger a priori nicht kennen.

Protocol Independent Multicast (2)

- [Arbeitsweise PIM-SM](#) (nutzt Reverse Path Forwarding): Phase 1
M-Empfänger trägt sich in M-Group G ein. Im lokalen Subnetz des Empfängers fungiert ein PIM-Router als DR (Designated Router). Dieser erzeugt Wildcard-Join-PDU in der Form (*,G), wird geroutet zu einem PIM-Router der Gruppe oder zu ausgezeichnetem Router, genannt Rendezvous Point RP. RP ist Vermittler zwischen M-Sender und Empfänger einer Gruppe und dient als Wurzel des M-Baumes (RP Tree).
Jeder (*,G) empfangene Router gehört dann zu RP Tree.
Sender sendet IP-Pakete zu seinem DR, der IP-Multicasts in Unicast IP einschlägt und als PIM-Registerpakete an RP schickt.
Von dort Multicast über RP Tree

Protocol Independent Multicast (3)

□ Arbeitsweise PIM-SM: Phase 2 (Register Stop)

Multicast zwischen Sender-DR und RP zunächst als encapsulated IP-unicast. Nach Empfang des ersten Pakets Überprüfung im RP, ob nicht statt Encapsulation besser native Multicast. Hierzu schickt RP spezielle Join-PDU (S,G) an Sender. PIM-Router auf diesem Rückweg, die bereits zu (*,G) gehören, können sich beim Sender-DR registrieren und somit Unicast-Pfad verkürzen.

Falls native Multicast-Pakete von neu registriertem PIM-Router eintreffen, verwirft RP eingepackte Multicasts durch Register Stop PDU an DR. Zu diesem Zeitpunkt existieren zwei Bäume: RP-Tree von RP an Empfänger, und Source Specific Tree vom Multicast Sender an den RP. Schnittpunkt ist Übergang von Unicast zu Multicast.

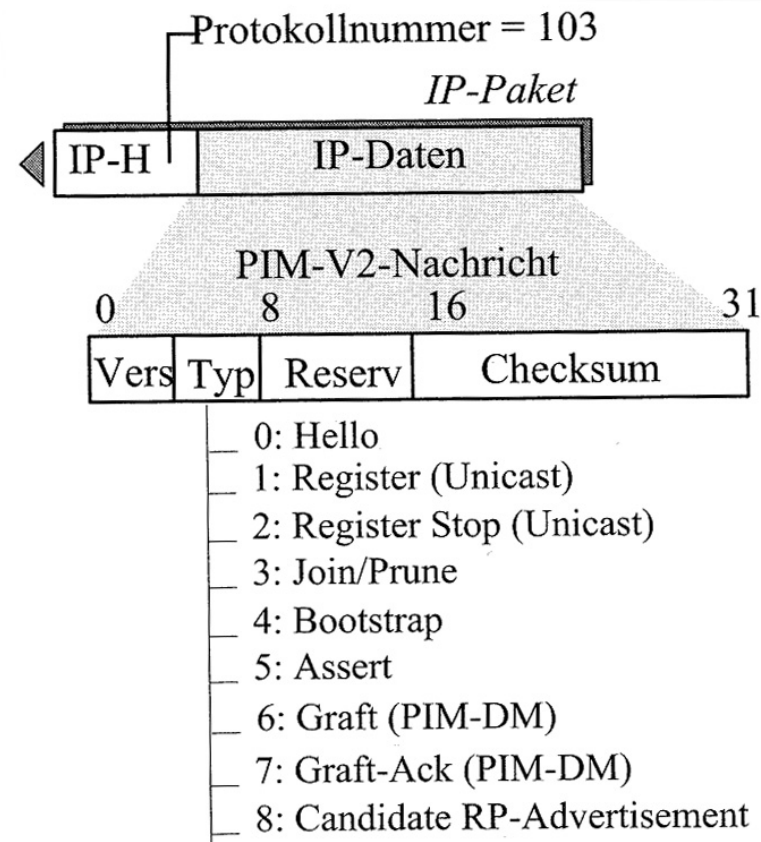
Protocol Independent Multicast (4)

□ Arbeitsweise PIM-SM: Phase 3 (Shortest Path Tree)

Weitere Optimierung durch Auswahl kürzester Pfad zwischen R und S. Dazu generiert Empfänger DR von sich aus Join PDU (S,G) in Richtung Sender. Damit wird empfängerspezifischer Shortest Path Tree SPT aufgebaut. M-Pakete kommen nur über SPT und RP-Tree. Letztere werden verworfen mittels Prune PDU (S,G,rpt) in Richtung RP. Prune-Nachricht kommt bei RP an oder beim letzten PIM-Router, der noch über RPT versorgt werden muss. Somit ist SPT etabliert.

Protocol Independent Multicast (5)

□ PIM-Nachrichten



Core Based Trees CBT

- ❑ Ähnliche Prinzipien wie bei PIM, RFC2189/2201
 - Core Router vergleichbar mit Rendezvous Point bei PIM
 - Empfänger trägt sich in M-Group ein durch IGMP-PDU an lokalen Router
 - Erreicht Nachricht einen CBT-Router, der den Core Router kennt, wird JOIN-PDU in Richtung Core Router gesandt.
 - Jeder Router auf dem Weg merkt sich den Forwarding State und sendet Bestätigung an den vorgehenden Router
- ❑ CBT-Nachrichten: HELLO, JOIN, QUIT-Notification, ECHO, FLUSH_TREE