

# Multicasting

Vortrag im Rahmen der Netzwerkausbildung am RRZE

26.2.1997

Toerless Eckert

---

Universität Erlangen-Nürnberg  
Martensstraße 1  
D-91058 Erlangen  
tte@bhn.de

folien.doc.recover: 26.02.97 12:35

## Inhaltsübersicht

- Grundlagen des Multicasting
  - ⇨ *spanning-trees, RPF, pruning, CBT, sparse, dense*
- IP-Multicasting
  - ⇨ IGMP, DVMRP, MOSPF, PIM, Tunnel
- Multicasting und Bridges
  - ⇨ IGMP-Snooping, GARP, GMRP
- ATM und Multicasting
  - ⇨ PTMP-VC, BUS, MARS
- Ausblick

folien.doc.recover: 26.02.97 12:35

## Was ist Multicasting ?

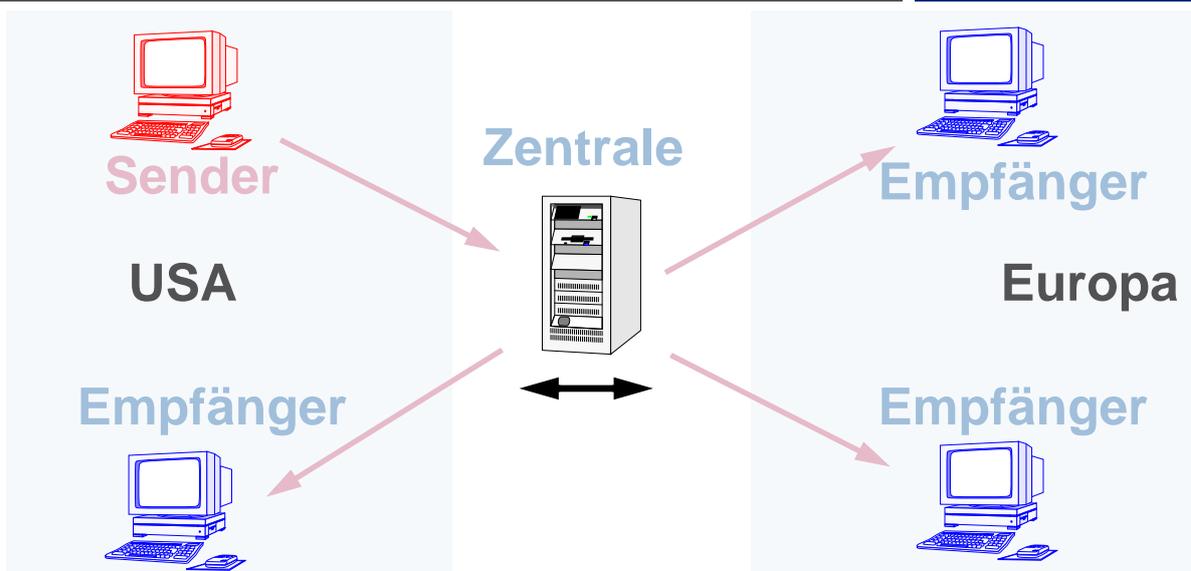
### Grundlagen

- **Unicasting:** Übertragung von Daten von einer Quelle zu einem Ziel, Punkt-zu-Punkt-Kommunikation, etc...
- **Multicasting:** Erster Ansatz: Übertragung von Daten von einer Quelle gleichzeitig an mehrere Ziele. Punkt-zu-Mehrpunkt-Kommunikation.
  - ⇒ Effizienz ?
- **Broadcast:** Ausnutzung spezieller Eigenschaften des Übertragungsmediums (z.B: Funk, LANs).
- Sternförmige Datenverteilung über eine Zentrale ??
- Allgemeinere Mechanismen zur Erzielung von Effizienz ?

Multicasting in einem Datennetz bezeichnet die quasi-simultane Übertragung von Daten einer Quelle an mehrere Ziele (eine Gruppe) unter besonderer Berücksichtigung der Effizienz dieser Übertragung im genutzten Datennetz => **Gruppenkommunikation.**

## Sternförmige Kommunikation

### Grundlagen

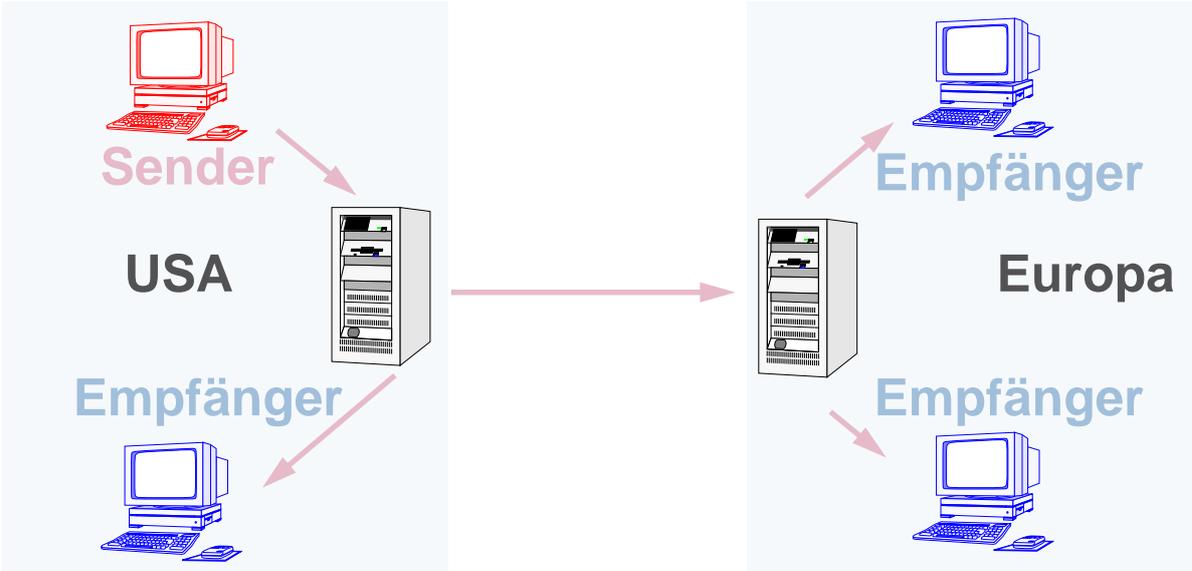


Ziel: Bezahlen Sie nur einmal Leitungskosten USA-Europa.  
Mit rein sternförmiger Verteilung ?

⇒ **Das Funktioniert Nicht**

# Baumförmige Kommunikation

## Grundlagen

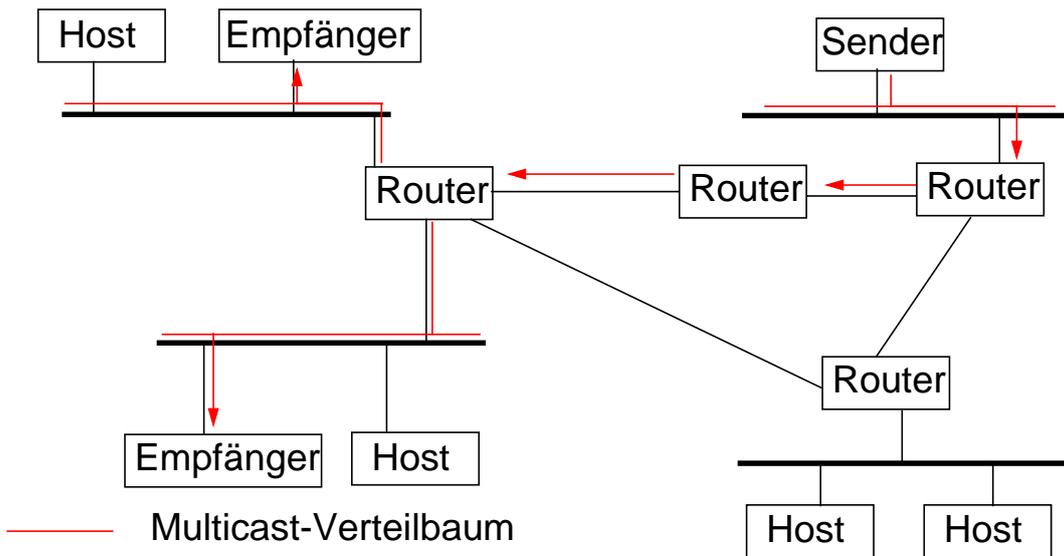


**Das Netz ist ein Wald, und deswegen müssen die Daten baumförmig verteilt werden !**

folien.docx.recover: 26.02.97 12:35

# Beispiel: Multicastverteilung

## Grundlagen



**Übertragung der Multicastdaten im Netz auf einem Baum:**

- Was für eine Baumart (*routing algorithm*) ?
- Wie den Baum auf die Empfänger zurückschneiden (*pruning*) ?

folien.docx.recover: 26.02.97 12:35

## Multicastbäume (1)

### Grundlagen

#### Gefluteter 'Baum':

- Paket auf allen Schnittstellen (außer der Empfangsschnittstelle) abschicken, außer es ist ein "doppeltes Paket".
- ⇒ Funktioniert prima im Usenet, bei Multicasting aber illusorisch, weil Erkennung doppelt empfangener Pakete im allgemeinen zu ressourcenaufwendig ist.

#### Steiner-Bäume:

- Der Baum, bei dem die Summe der Kosten aller Netzverbindungen, die den Multicastbaum bilden, minimal ist.
  - ⇒ Berechnung NP-vollständig (über die Teilnehmer des Multicast). Sehr instabil bei Änderung der Teilnehmer.
  - ⇒ Theoretisch interessant, aber nicht verwendbar im realen Netz.

## Multicastbäume (2)

### Grundlagen

#### (global) aufgespannter Baum:

- ⇒ Aktive Pfade im Netzwerk auf einen aufgespannten Baum reduzieren, auf diesen Pfaden fluten.
- Einfach, aber ineffizient in redundanter Topologie. Konzentrierung des Verkehrs auf wenige Pfade im Zentrum => Überlastung.

Verbesserung durch Definition eines **Kerns**  
(CBT - **Core Based Trees**):

- **Kern:** Feste Adresse pro Multicastgruppe.
- Empfänger einer Gruppe schicken "Anmeldungen" an den Kern. Router auf dem Weg bauen daraus den Baum auf:
  - ⇒ Schnittstelle, von der die Anmeldung kam, in den Baum einfügen.
  - ⇒ Anmeldung nur weiterleiten, wenn es die erste der Multicastgruppe auf dem Router ist.

## Multicastbäume (3)

### Grundlagen

#### Am Sender aufgespannte Bäume:

- Verteilte Berechnung des aufgespannten Baumes in Abhängigkeit des Absenders durch **Reverse Path Forwarding** (RPF):
  - ⇨ Jeder Router leitet Pakete vom Absender genau dann an alle seine Schnittstellen (außer der Empfangsschnittstelle) weiter, wenn die Empfangsschnittstelle diejenige ist, auf der der Router Unicastpakete an den Absender schicken würde.
  - ⇨ Verfeinerungen möglich, wenn lokaler Router etwas über die Routingtabellen der nachfolgenden Router weiß.
  - ⇨ Einfach zu berechnen.
  - ⇨ Größere Auslastung der Netzkapazitäten bei mehreren Sendern.
  - ⇨ Immer der kürzeste Pfad von Absender zu den Empfängern (**Shortest Path Tree Forwarding - SPF / SPT**).

## Multicastbäume (4)

### Grundlagen

#### Pruning bei SPT (*flood & prune*):

- Prinzip "Fluten und zurückschneiden". Erstes Paket fluten, woraufhin von den Blättern, an denen das Paket ankommt ohne gebraucht zu werden, "**Prune-Nachrichten**" zurückgeschickt werden. Diese Pakete etablieren Zustand in den Routern, daß ab hier der Teilbaum nicht mit Paketen versorgt wird.
  - ⇨ Zyklisch wegen Variabilität der Gruppenzugehörigkeit.

#### Zwei grundsätzliche Möglichkeiten eines Routingprotokolls, Multicastverkehr zuzustellen:

- **dense-mode**: Verkehr wird überallhin zugestellt, wer ihn nicht haben will, muß ihn explizit abbestellen.
- **sparse-mode**: Verkehr wird nirgends zugestellt, sondern muß explizit beantragt werden.

- **IP-Multicasting:**

IP-Pakete verschicken an eine Gruppe von Empfängern:

- ☞ Effizienter Versand an die Empfänger. Ziel: Genau eine Kopie eines Paketes überquert genau die Teilnetze, die auf dem Weg zu Gruppenmitgliedern liegen.

- **Adressierung:** (IP-Multicast-Gruppenadressen)

IP Class-D Adressen (224.0.0.0 bis 239.255.255.255).

- **API (Socket/Unix):**

- Standard BSD-Socket Schnittstelle mit Gruppenadressen.
- Nur unterstützt bei UDP (nicht TCP).
- Wenige neue Socketoptionen:
  - IP\_ADD\_MEMBERSHIP, IP\_DROP\_MEMBERSHIP,
  - IP\_MULTICAST\_TTL.

### **Host support for IP Multicasting (RFC1112)**

- **IGMP** - Das *Internet Group Management Protocol* ist eine Ergänzung des IP-Protokolls (wie ICMP), mit dem Rechner aktiv auf dem Netz ankündigen, wenn sie als Mitglied einer Multicastgruppe geführt werden wollen.
  - ☞ Ankündigungen werden regelmäßig wiederholt. Wenn Ankündigungen anderer Gruppenmitglieder gehört werden, so werden die Eigenen unterdrückt.
- **Abbildung von IP-Multicastadressen auf MAC-Adressen:**

In IEEE802-MAC Netzen werden IP-Multicastadressen in einen 23-Bit Bereich der MAC-Multicastadressen (01-00-5E-XX-XX-XX) gemappt.

  - ☞ Verwendung von MAC-Multicastadressen bewirkt effektive Filterung in den LAN-Schnittstellen der Endsysteme am Netz.

- **IGMP** - Damit stellt der Router die Existenz von Gruppenmitglieder auf seinen direkt angeschlossenen Netzen fest.
- **IP-Multicast-Routingprotokoll:** Zur Berechnung der im Router verwendeten Weiterleitungstabellen für IP-Multicastverkehr.
  - ⇨ DVMRP, MOSPF, CBT, PIM {sparse, dense}
- **Weiterleitung von IP-Multicastverkehr:** (im Gegensatz zu Unicastverkehr)
  - ⇨ Packet verwerfen, z.B.: aufgrund eines “**RPF-checks**”.
  - ⇨ Ansonsten Paket ggf. duplizieren und auf die durch die Weiterleitungstabelle festgelegten Schnittstellen legen.
  - ⇨ Paket an der Schnittstelle ggf. verwerfen wegen “**TTL-Threshold-Check**”, wenn TTL des Pakets nicht größer ist als dieser Threshold der Schnittstelle (geographische Grenze).

- **DVMRP:** *Distance Vector Multicast Routing Protocol* (RFC 1075)  
**Das Standard-Multicast-Routingprotokoll, weil am ältesten.**  
 SPT-Weiterleitung mit **flood&prune**. Verwendet Routingtabellen von eigenem RIP(2)-artigen Routingprotokoll nicht nur für RPF-Checks, sondern auch für weitergehende Verkehrsreduzierung. Ergänzung z.B.: **graft**-Nachrichten, um existierende **prunes** bei neuen Empfängern rückgängig zu machen.
  - ⇨ DVMRP ist ein **dense-mode** Protokoll.
- **MOSPF:** Ergänzung von OSPF um Multicasting (RFC 1584)  
 IP-Multicastgruppenzugehörigkeit wird in der LS-Datenbank verteilt, so daß jeder Router ohne weitere “Erkundigungen” auf dem Netz die Verteilungsbäume berechnen kann.
  - ⇨ **MOSPF ist ein sehr gutes sparse-mode Protokoll mit potentiellen Skalierungsproblemen.**

## Routingprotokolle (2)

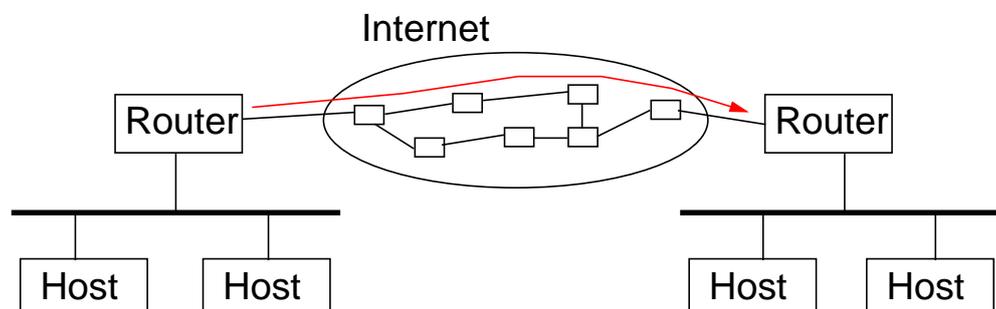
### IP-Multicasting

- **PIM: Protokoll Independent Multicasting**  
PIM verwendet nur die Unicast-Routingtabellen im Gegensatz zu DVMRP oder MOSPF.
  - ⇒ PIM erfordert immer strikt symmetrische Routen.
- **Dense-mode PIM:**  
Ähnlich wie DVMRP, aber viel einfacher.
- **Sparse-mode PIM:**  
Im Prinzip wie CBT mit Ergänzungen. Der **Kern** heißt **Relay-Point** und nachdem Verkehr vom Sender über den CBT an Empfänger geschickt wird, kann jeder Router (z.B.: auf Grund der Last) entscheiden, den Verkehr über einen SPT zu den Empfängern zu routen. Sehr komplex, aber gut geeignet für große Netze.

## Tunnel

### IP-Multicasting

**Virtuelle Verbindungen zwischen entfernten Routern, um dazwischenliegende Router zu umgehen, die nicht multicast-fähig sind.**



- Verbreiteter Mechanismus, nicht allein bei IP-Multicasting.
- Diese bei IP-Multicasting sogenannten DVMRP-Tunnel arbeiten mit *IP-in-IP*-Verpackung (Protokoll 0x04).

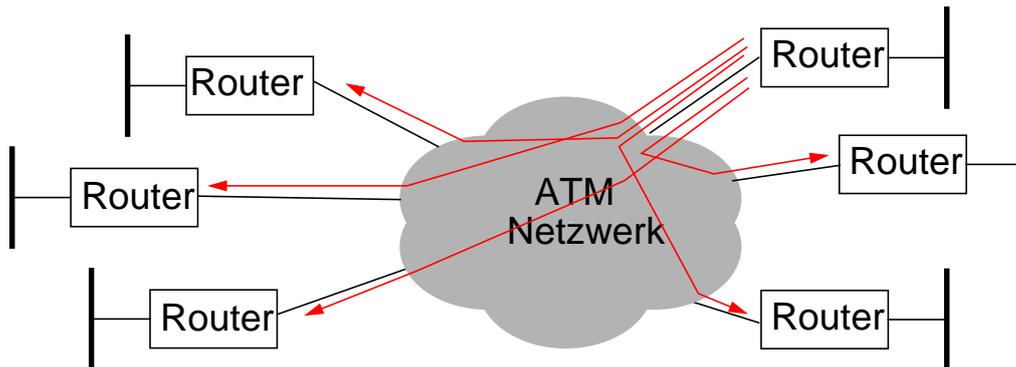
## Multicast auf Backbone-Netzen

### IP-Multicasting

Backbone-Netze sind oft Non-Broadcast-Multi-Access (NBMA) Netzwerke ohne ausreichende Multicast-Fähigkeiten:

- ATM, Frame-Relay, SMDS, X.25
- Multicastverkehr muß durch Tunnel transportiert werden, aber dies skaliert nicht!

⇒ **Lösung: Multicasting des unterliegenden Netzes ausnutzen (ATM).**



## Einführung

### Multicasting und Bridges

- ⇒ Multicasting bei LAN/MAN mit IEEE 802 MAC ist definiert. Nutzung durch: IP, IPX, Appletalk und andere Protokolle.
- ⇒ In lokalen Netzen werden Bridges (Switches) verwendet, um die Kapazität zu vergrößern und verschiedene Netztypen zu verbinden: FDDI, 10 Mbps Ether, 100 Mbps Ether, Token Bus,...
- ⇒ Bridges nach IEEE 802.1D (Spanning Tree) kennen derzeit keine Mechanismen, um Multicastverkehr zu filtern.

**Multicastpakete sind in allen Segmenten eines gebriidgten LAN wie Broadcastpakete sichtbar. Folge: unnötige Last.**

**Lösungsansatz: Intelligere filternde Bridges:**

Jetzt: **IGMP-Snooping** (nur IP)

Später: **GARP/GMRP** (alle Protokolle)

## IGMP-Snooping

### Multicasting und Bridges

Bridge erkennt (*snoops*) IGMP-Pakete und leitet daraufhin die Pakete passend nur auf die Segmente mit Multicastteilnehmern.

- ⇒ Cisco / Catalyst 1200: IGMP-Snooping
- ⇒ 3Com / Lanplex-Series: Autocast-VLANs

#### Probleme:

- Nur für IP-Multicastverkehr geeignet
- Viele Switches können IGMP-Pakete nicht von IP-Multicast-Datenpaketen unterscheiden (gleiche IP-Adressen).
  - ⇒ IGMP-Snooping so nicht realisierbar.
  - ⇒ Hilfsprotokoll für "dumme" Bridges/Switches wird benötigt:
    - Cisco: Catalyst 5000 / 3x00 / 2xxx Switches CGMP  
Router am LAN wandelt IGMP-Pakete um in Steuerpakete für die Bridges (Switches).

## GARP/GMRP (1)

### Multicasting und Bridges

IEEE802.1p: "Traffic Class Expedition and Dynamic Multicast Filtering"

⇒ Integrierter Ansatz für die Filterung von Multicastverkehr in IEEE 802-MAC LANs:

- Für Rechner und Bridges.
- Registrierung von Gruppenmitgliedschaft (wie IGMP) und Aufbau des Verteilungsbaums (wie IP-Routingprotokolle).
- Einfache Verteilung im gebridgten LAN per **Spanning-Tree**.
- Ermöglicht **Sender-Pruning**.
- Bietet Kompatibilität mit Komponenten ohne GARP/GMRP.
- **Sparse**-Verhalten (wenn alle Komponenten GMRP können).

⇒ Derzeit in früher Phase der Standardisierung (**5th draft**).  
(Pre-)CGMP wird Mitte 1997 auf PCs und LAN-Switches verfügbar.

## GARP/GMRP (2)

### Multicasting und Bridges

### GARP: *Generic Attribute Registration Protocol*

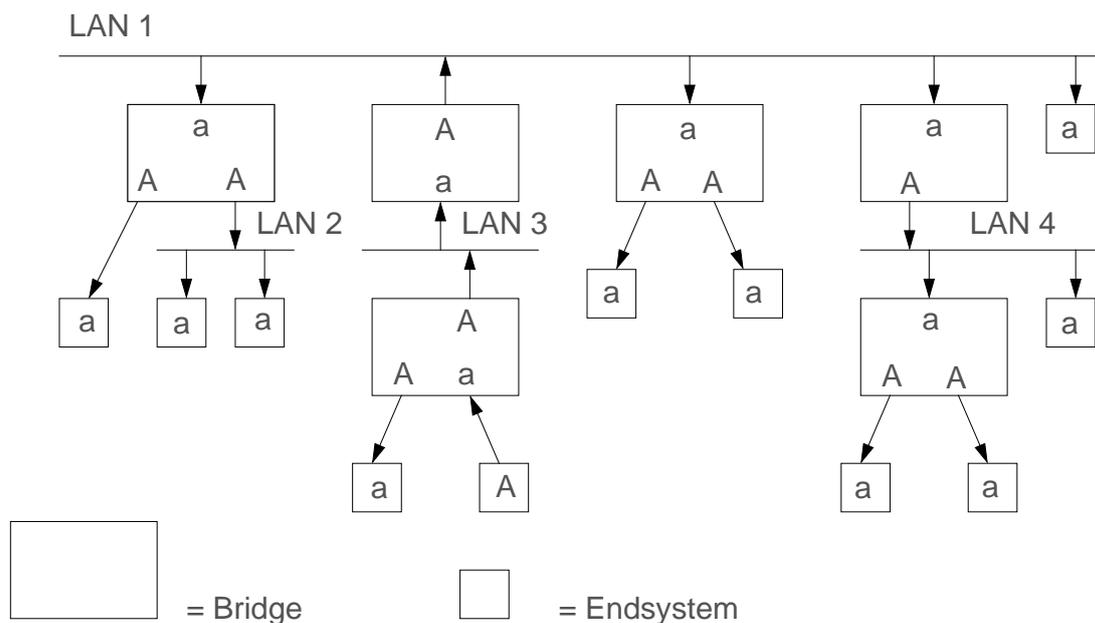
- Allgemeines Protokoll im LAN zur Verteilung von Attributen zwischen Teilgruppen von Rechnern und Bridges im LAN.

### GMRP: *GARP Multicast Registration Protocol*

- Protokoll zur Etablierung von zwei Attributen im LAN unter Nutzung von GARP zur Verteilung:
  - Teilnahme von Rechnern und Bridge-Ports an Multicastgruppen
  - Art der Filterung auf Bridge-Ports :
    - A) Alle Multicastpakete auf den Port weiterleiten.
    - B) Alle nicht im LAN registrierten Adressen weiterleiten (nur registrierte Adressen filtern).
    - C) Alle nicht im LAN registrierten Adressen filtern.

## GARP/GMRP (3)

### Multicasting und Bridges



→ Richtung der Verteilung des Attributes A

A = Deklaration des Attributes A

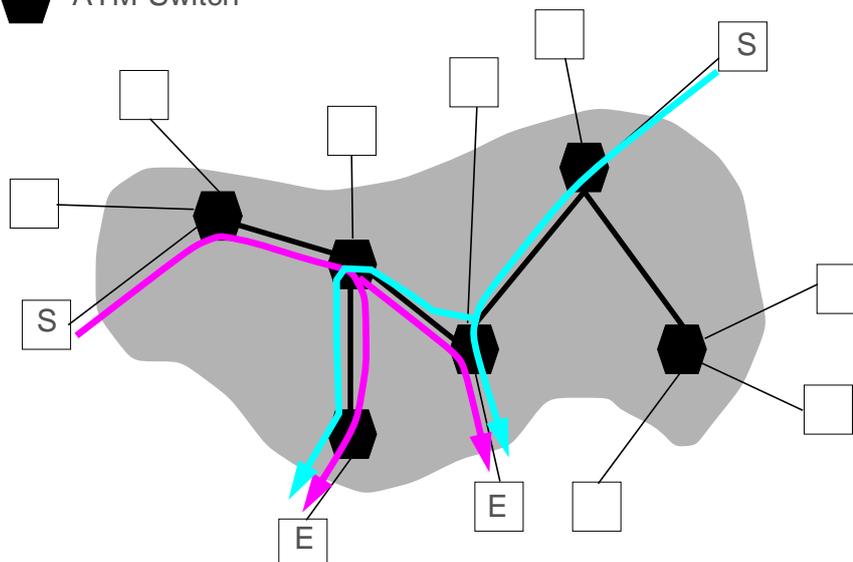
a = Registrierung des Attributes A

Multicast bei **Native-ATM**: *Point to MultiPoint Virtual Channels*:

- Virtuelle Kanäle bei ATM sind unidirektional.
- PTMP-VC sind Verteilbäume vom Sender zu den Empfängern (durch beliebig viele ATM-Switches).
  - ▷ Für jeden Sender eigener PTMP-VC notwendig:  
ATM-Zellen haben keine explizite Absenderadresse,  
Zellen mehrerer Sender würden sich bei "MPTMP" vermischen.
- ATM-Standard: Jeder (heutige) ATM-Switch unterstützt sie.
- Konfiguration:
  - ▷ statisch durch Konfiguration in den Switches (PTMP-PVC).
  - ▷ dynamisch durch Signalisierung vom Sender (PTMP-SVC)  
"add <dest> to <VC>", "del <dest> from <VC>".

## ATM PTMP-VC Beispiel

□ Endsysteme mit ATM-Verbindung S = Sender E = Empfänger  
 ● ATM-Switch



VC-Mesh

## IP-Multicasting

## ATM und Multicasting

Wird ATM als Transportnetz für IP verwendet, entscheidet die Art der Emulation eines IP-Subnetzes im ATM über den Multicastdienst:

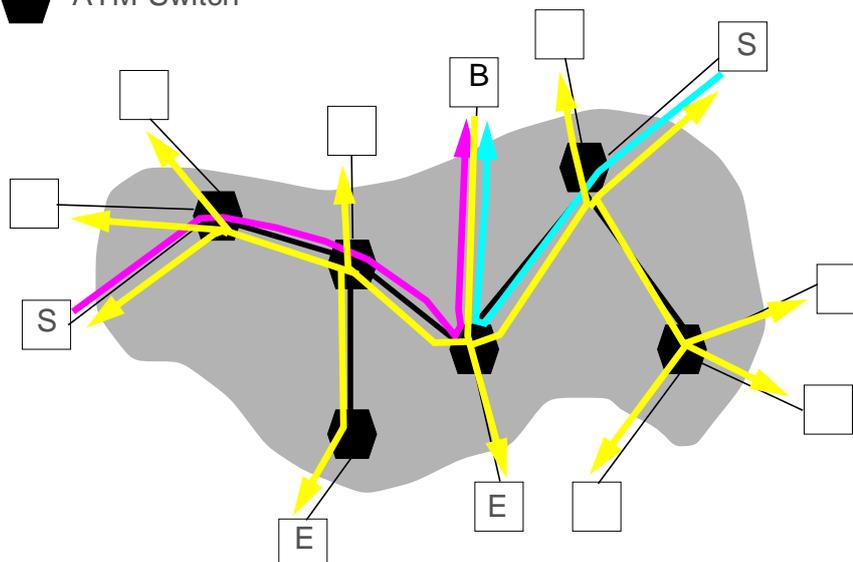
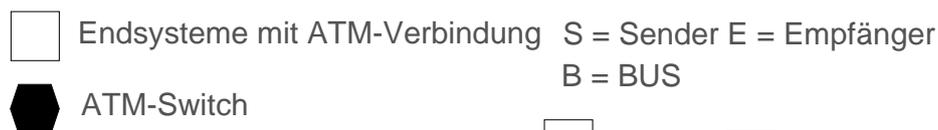
- Classical-IP (RFC1577) oder statische VC-Konfiguration:
  - ◊ Nur PTP-VC, kein Broadcast, kein Multicast.
- LAN-Emulation
  - ◊ Multicastverkehr ist wie Broadcastverkehr (wie im LAN !!)
  - ◊ Broadcastverkehr läuft vom Sender über PTP-VC zum BUS (*Broadcast and Unknown Server*) und von dort über einen PTMP-VC zu allen Teilnehmern des VLAN.

### Folgen:

Je nach Plazierung des BUS durchläuft Multicastverkehr zweimal das ATM-Netz. Alle Stationen am E-LAN müssen den Verkehr empfangen. Der BUS wird leicht zum Engpaß.

## LAN-E Multicastverkehr

## ATM und Multicasting



█ Verkehr vom BUS  
 (Summe der ankommenden Daten am BUS)

## Broadcast-Server

## Besseres IP-Multicasting ?

ATM und Multicasting

- Proprietäre Lösungen:
  - Fore-IP. Nahezu 1:1 Umsetzung der ATM-Fähigkeiten auf IP. Effizientes IP-Multicasting.
  - IP-Switching ? (RFC 1953, ..) ?

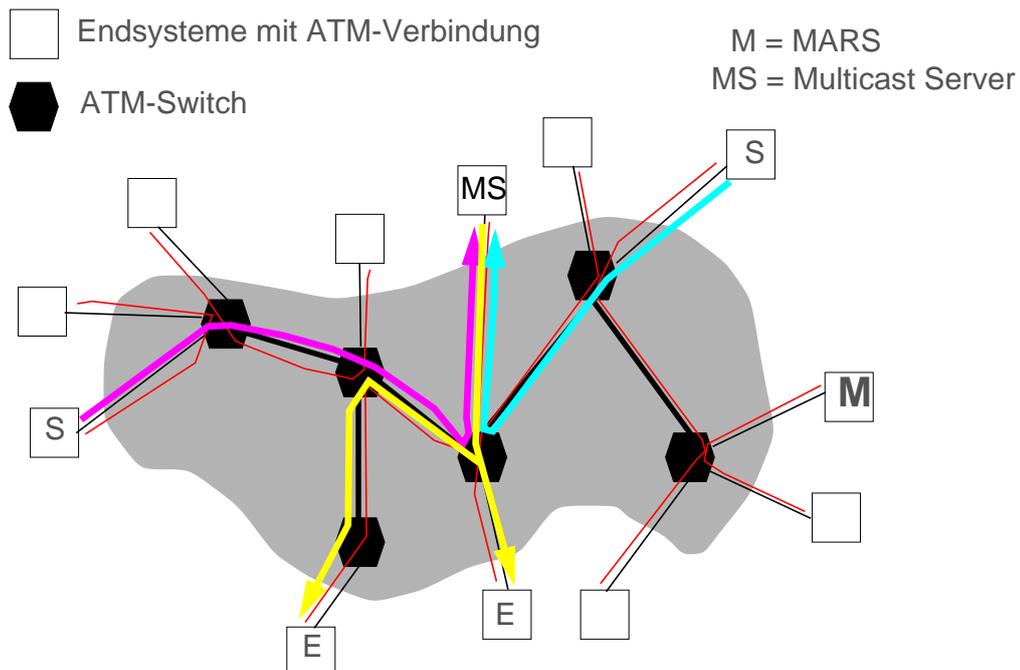
...

Proprietäre Lösungen funktionieren JETZT in HOMOGENEN Netzen

- Standardlösungen:
  - LAN-E + GARP/GMRP +++ ???
  - Classical-IP oder MPOA (Multiprotokoll over ATM) + **MARS: Multicast Address Resolution Server (RFC2022):**
    - Server im Netz (MARS) verwaltet Gruppenzugehörigkeit, Multicastverkehrsweiterleitung jedoch durch: **VC-Meshes** oder **Multicast-Servers**

## MARS: Multicast Adress Resolution Server

ATM und Multicasting



Multicast-Server

## Ausblick

### Stand IP-Multicast-Installation:

- ⇒ **MBone:** Verbund der Einrichtungen, die weltweit miteinander verbunden IP-Multicast betreiben. Seit 1992. Erlangen seit 1993.
- ⇒ **MBone in Deutschland:** derzeit ca. die Hälfte aller Universitäten im MBone erreichbar, in den meisten Fällen jedoch nur jeweils kleine Inseln.
- ⇒ **MBone in Bayern:** Alle Universitäten sind angeschlossen, bis auf Erlangen und München jedoch auch nur in kleinen Inseln.
- ⇒ **Erlangen:** Alle Backbone-Router haben IP-Multicasting konfiguriert. So kein weiterer Router (z.B.: Novell) zwischen einem Backbonerouter und dem Rechner ist, funktioniert Multicast.
- ⇒ Fragen, Information, Mailverteiler:  
[help@mbone.de](mailto:help@mbone.de) / [www.mbone.de](http://www.mbone.de) / [mbone-de-request@mbone.de](mailto:mbone-de-request@mbone.de)