IBM

Dr. Torsten Braun

European Networking Center

Heidelberg

IP Multicast

Oktober 1996 - 1

IP Multicast

Dr. Torsten Braun

Tel.: 06221/59-4352
Email: braun@heidelbg.ibm.com

Meet The Experts
Internet - Technik und Anwendung

8. Oktober 1996

IBM European Networking Center

IBM

Dr. Torsten Braun

European Networking Center

Heidelberg

IP Multicast

Oktober 1996 - 2

© IBM Deutschland 1996

Inhalt

- Einführung
- Internet Group Management Protocol (IGMP)
- IP Multicast in LANs und ATM
- Multicast Routing
  - Grundlegende Verfahren
  - Distance Vector Multicast Routing Protocol (DVMRP)
  - PIM (Protocol Independent Multicast)
  - MOSPF (Multicast Open Shortest Path First)
- Multicast Backbone (Mbone)
  - Topologie
  - Anwendungen

Dr. Torsten Braun

European Networking Center

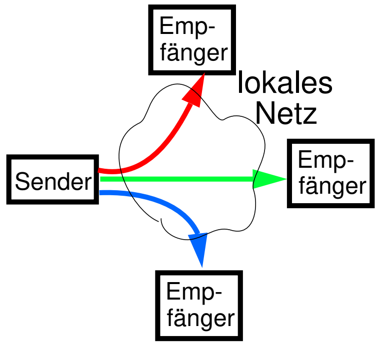
Heidelberg

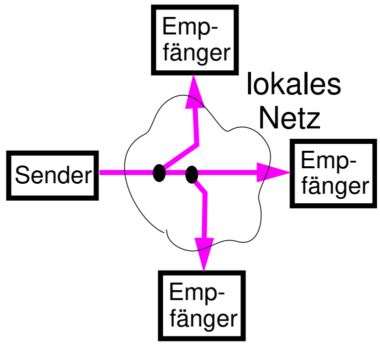
Multicast-Kommunikation

Unicast: 1:1-Kommunikation
Multicast: 1:n-Kommunikation

Vorteil von Multicast:
einmaliges Senden eines Pakets für
n Empfänger statt
n-faches Senden eines Pakets
bei Unicast

-> geringere Belastung des Senders
-> geringere Netzbelastung





Dr. Torsten Braun

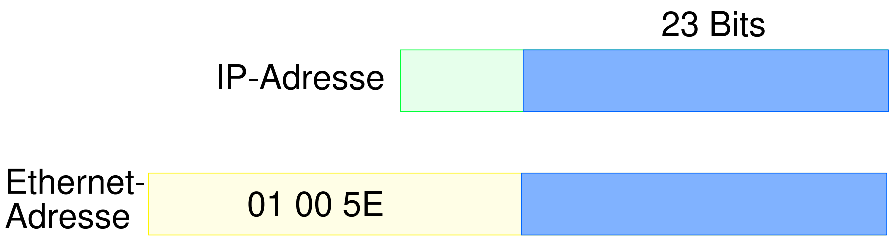
European Networking Center

Heidelberg

© IBM Deutschland 1996

Abbildung von IP-Multicast-Adressen auf LAN-MAC-Adressen

- keine Adreßauflösung in IEEE 802 LANs für IP-Multicast-Adressen
- automatische Ableitung der MAC-Adresse aus IP-Multicast-Adresse (Beispiel: Ethernet)
  - Broadcast-Adresse (0x FF FF FF FF FF FF)
  - Multicast-Adresse



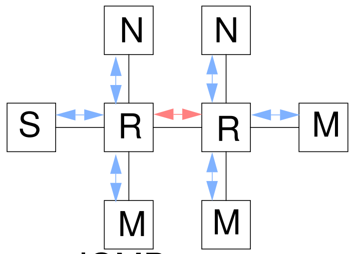
Dr. Torsten Braun

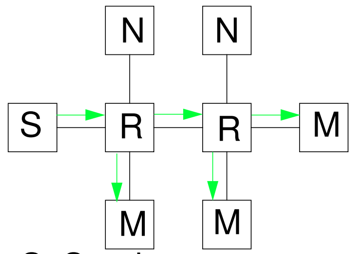
European Networking Center

Heidelberg

IP-Multicast-Modell

- Sender sendet Multicast-Paket an durch eine IP-Multicast-Adresse gekennzeichnete Gruppe.
- Sender kennt die Empfänger nicht.
- Jedes Endsystem kann einer Gruppe beitreten.
- UDP unterstützt Multicast, TCP jedoch nicht
  => Multicast-Kommunikation ist Datagramm-orientiert
- Aufbau eines Multicast-Baums





IGMP
Multicast-Routing-Protokoll
Daten

S: Sender
M: Mitglied der Multicast-Gruppe
N: Nichtmitglied
R: Router

Dr. Torsten Braun

European Networking Center

Heidelberg

Multicast-Adressierung

- Multicast-Adressierung in IPv4:

4 bit	28 bit
1110	Gruppen-ID

- Beispiele vorreservierter Adressen:
  - 224.0.0.1 : all hosts
  - 224.0.0.2 : all routers
  - 224.0.0.4 : DVMRP routers

- Multicast-Adressierung in IPv6:

8bit	4bit	4 bit	112 bit
11111111	flags	scope	Gruppen-ID

Dr. Torsten Braun

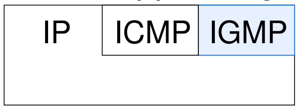
European Networking Center

Heidelberg

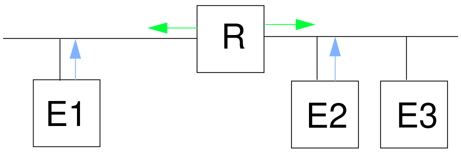
Gruppenverwaltung: IGMPv1

Router müssen wissen, ob lokale Gruppenmitglieder existieren

-> IGMP



- 2 Nachrichtentypen:
  - Membership Query (periodisch (z.B. 1/min) durch Router an die “all hosts”-Gruppe)
  - Membership Report (Antwort eines Hosts zur Anzeige der Gruppenzugehörigkeit)



E3 hört Antwort von E2 und unterdrückt eigene Antwort

- Dienstprimitive:
  - JoinHostGroup (address, interface) -> Senden v. Reports
  - LeaveHostGroup (address, interface) -> kein Senden v. Reports

Dr. Torsten Braun

European Networking Center

Heidelberg

Gruppenverwaltung: IGMPv2

- Problem von IGMPv1:
  große Verzögerung beim Feststellen der Gruppenzugehörigkeiten wegen der periodischen Membership Query Paketen
- 2 Arten von Membership Query Paketen:
  - general
  - group-specific
- zusätzliches Paket: Leave Group
  - Endsystem setzt Flag, falls es auf das letzte Membership Query geantwortet hat.
  - Endsystem sendet beim Verlassen der Gruppe ein Leave Group, falls Flag gesetzt ist.
  - Router sendet anschließend ein group-specific Membership Query.
    -> Router stellt schneller fest, wenn keine weiteren Gruppenmitglieder existieren

IBM

Dr. Torsten Braun

European Networking Center

Heidelberg

IP-Multicast in ATM

- ATM kennt derzeit (ATM Forum UNI3.x) keine Multicast-Adressen
- ATM ist verbindungsorientiert
- ATM-Multicast-Unterstützung durch Punkt-zu-Mehrpunkt-Verbindungen (Multipoint-VCs)
- Problem:
  Abbildung von IP-Multicast-Paketen auf ATM-Multipoint-VCs
- Lösungsansätze:
  - LAN Emulation
  - Multicast-Router als Multicast-Server
  - Multicast Address Resolution Server (MARS)

IBM

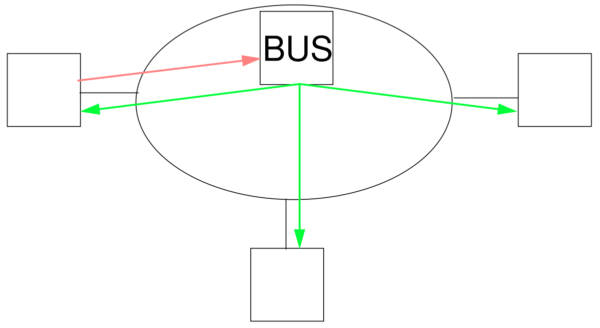
Dr. Torsten Braun

European Networking Center

Heidelberg

LAN Emulation

- Funktionalität eines IEEE 802 LANs wird durch
  LAN Emulation Service emuliert.
  Multicast-Pakete werden an BUS (Broadcast and unknown server) gesendet.
  BUS verteilt Multicast-Pakete per Broadcast.



- Nachteile:
  - Broadcast der Multicast-Pakete
  - Paketgröße auf maximale Größe in Ethernet oder Token Ring begrenzt

IBM

Dr. Torsten Braun

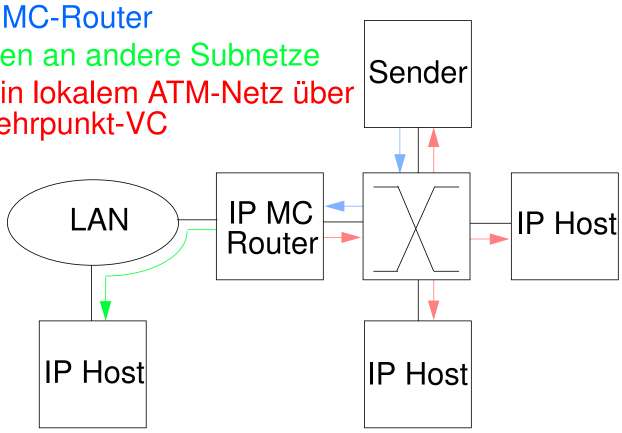
European Networking Center

Heidelberg

Multicast-Server

- Multipoint-VCs basierend auf Analyse von IGMP-Paketen

1. Paket an MC-Router
2. Weiterleiten an andere Subnetze
3. Verteilen in lokalem ATM-Netz über Punkt-zu-Mehrpunkt-VC



- Vorteile: kein zusätzliches Protokoll, geringe Änderungen an Endsystemimplementierungen
- Nachteile: Router ist Engpaß, zusätzlicher Hop für lokalen Verkehr

IBM

Dr. Torsten Braun

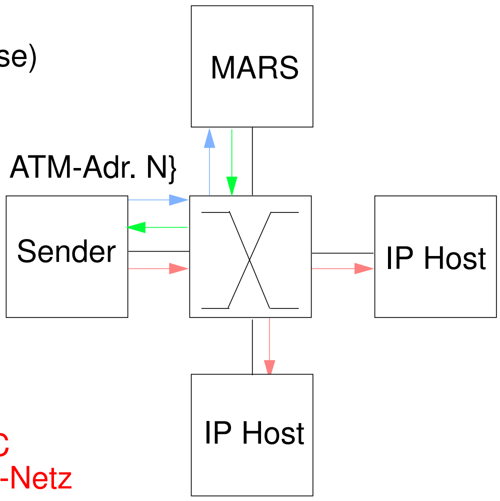
European Networking Center

Heidelberg

MARS: VC Mesh

ATM/IP-Endsysteme senden MARS_JOIN(ATM, IP-MC-Adresse) zum Beitreten einer Gruppe

MARS verwaltet Listen {IP_MC-Adresse, ATM-Adr. 1, ..., ATM-Adr. N}



1. Adreß-Anfrage an MARS
2. Mitteilen der Adressen der einzelnen Gruppenmitglieder
3. Aufbau eines Multipoint-VC und Verteilen in lokalem ATM-Netz

- Vorteil: direkte Kommunikation zw. Sender u. Empfänger
- Nachteile: n·m Multipoint-VCs für n Sender und m Gruppen, hoher Signalisierungs-Aufwand zum Verteilen der Gruppeninformation

IBM

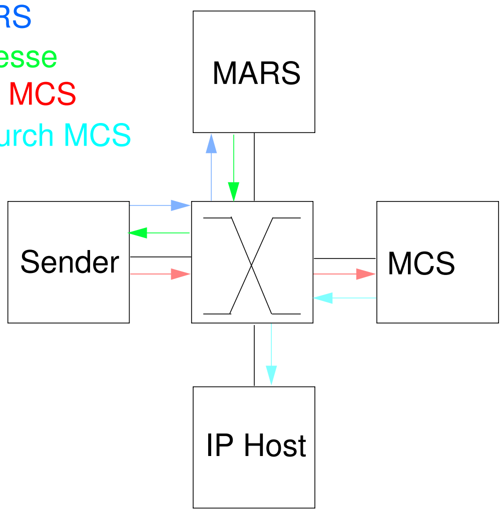
Dr. Torsten Braun

European Networking Center

Heidelberg

MARS: Multicast Server (MCS)

1. Adreß-Anfrage an MARS
2. Mitteilen der MCS-Adresse
3. Senden des Pakets an MCS
4. Verteilen des Pakets durch MCS



- Adreßauflösung identisch wie bei VC Mesh
- Mehrere MCSs müssen synchronisiert werden.

IBM

Dr. Torsten Braun

European Networking Center

Heidelberg

Multicast-Routing

- 2 grundsätzliche Verfahren
  - Distanz/Vektor-Routing

    Jeder Router besitzt Informationen über seine Entfernung zu jedem anderen Router und verteilt seine Routing-Informationen (z.B. Distanzen zu anderen Routern) an die benachbarten Router (Z.B. DVMRP)

  - Link-State-Routing

    Jeder Router besitzt Informationen über jeden Link der Domäne. Aus diesen Informationen können Distanzen zwischen allen Routern der Domäne berechnet werden (-> Berechnung der kompletten Topologie, z.B. MOSPF)

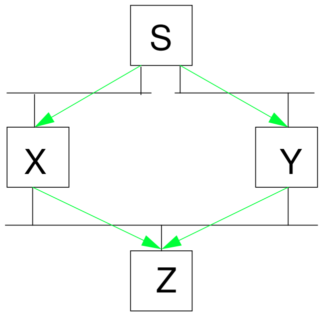
Dr. Torsten Braun

European Networking Center

Heidelberg

Reverse Path Flooding (RPF)

- Ein Paket wird auf allen Links weitergeleitet mit Ausnahme des Links, auf dem das Paket ankam.
- Bedingung: Das Paket muß auf dem kürzesten Weg angekommen sein, d.h. der Eingangslink muß den Ausgangslink für ein Paket zum Sender darstellen.



- Nachteile:
  - Broadcast-Verfahren
  - mehrfacher Empfang eines Pakets durch einen Host (Kollision)

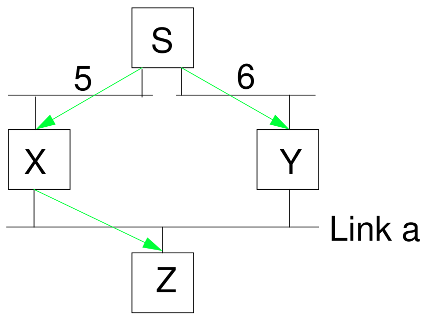
Dr. Torsten Braun

European Networking Center

Heidelberg

Reverse Path Broadcast (RPB)

- Vermeiden einer Kollision durch Definition einer Eltern-Kind-Beziehung
- X ist Elternteil eines Links, wenn seine Entfernung zur Quelle geringer als die von Y ist. Bei Entfernungsgleichheit gibt die kleinere Adresse den Ausschlag. Router tauschen Routing-Informationen untereinander aus, um Eltern-Kind-Beziehung festzustellen (Bsp.: Link a ist Kind von X, nicht von Y)
- Weiterleiten v. Paketen nur über Kind-Links -> spanning tree



- Nachteil: Broadcast-Verfahren

IBM

Dr. Torsten Braun

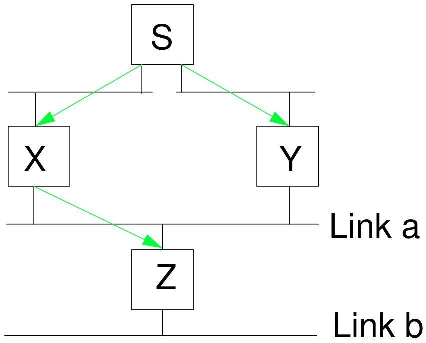
European Networking Center

Heidelberg

IP Multicast

Truncated RPB (TRPB)

- Erweiterung gegenüber RPB:
  - Pakete werden nicht auf Blatt-Links ohne Gruppenmitglieder weitergeleitet.
  - Erkennung von Blatt-Links durch Senden von Router-Nachrichten “Paket für Q wird über Link x gesendet”



- Bsp.: Z sendet Nachricht “Paket für S wird über Link a gesendet”
  -> Link a ist kein Blatt-Link
  -> Link b ist ein Blatt-Link

IBM

Dr. Torsten Braun

European Networking Center

Heidelberg

IP Multicast

© IBM Deutschland 1996

Reverse Path Multicast (RPM)

- Pakete werden zunächst wie bei TRPB gesendet.
- Einführung von Pruning:
  - Ein Router, dessen Kind-Links alle auch Blatt-Links ohne Gruppenmitglieder darstellen, sendet einen Non-Membership-Report (NMR) an den übergeordneten Router, d.h. an den vorhergehenden Router im Multicast-Baum.
  - Router, die von allen untergeordneten Routern NMRs empfangen haben, senden ebenfalls einen NMR an den übergeordneten Router.
  - NMRs enthalten eine Zeitangabe, nach der das Pruning wieder aufgehoben wird.
  - NMRs können auch aufgehoben werden, wenn ein Gruppenmitglied an einem Link aktiv wird (-> Analyse der IGMP-Nachrichten durch den Router).

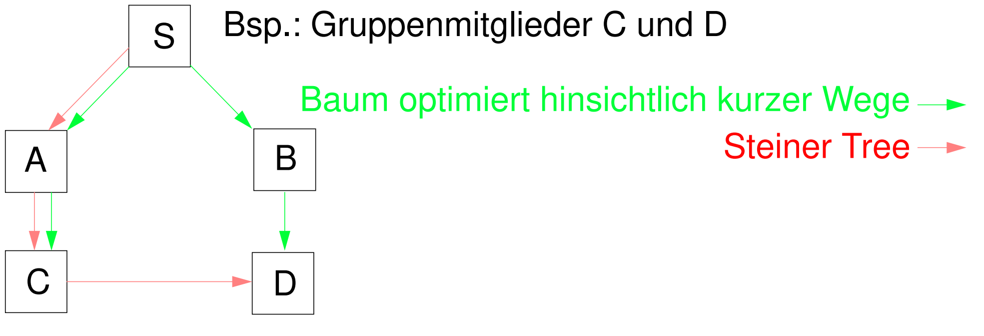
Distance Vector Multicast Routing Protocol (DVMRP)

- Erste Version von DVMRP (RFC 1075) basiert auf RPM
  - Routing-Information wird über DVMRP-Response-Nachrichten periodisch (1 min) oder bei Änderungen zwischen den DVMRP-Routern ausgetauscht.
  - DVMRP-Request-Nachricht zur Anforderung von Routing-Informationen bei Neustart eines Routers
- Austausch von Unicast-Routing-Information in DVMRP integriert
  - Vermeidung von Inkonsistenzen aufgrund unterschiedlicher Unicast-Routing-Protokolle
  - Benutzung unterschiedlicher Routen für Unicast- und Multicast-Verkehr möglich

IBM
Dr. Torsten Braun
European Networking Center
Heidelberg

Steiner Trees

- Vorteil von RPM/TPRB:
  optimierte Wege von einer Quelle zu den einzelnen Empfängern
- Nachteil: Erhöhter Verbrauch von Netzressourcen
- Abhilfe: Steiner Tree (Optimierung der Netzressourcen)



- Nachteil:
  Steiner Tree ist sehr instabil (z.B. A, C, D-> A, B, C, D)
  und aufwendig (NP-vollständig) zu berechnen
  -> Vereinfachung durch CBT-Ansatz

IBM

Dr. Torsten Braun

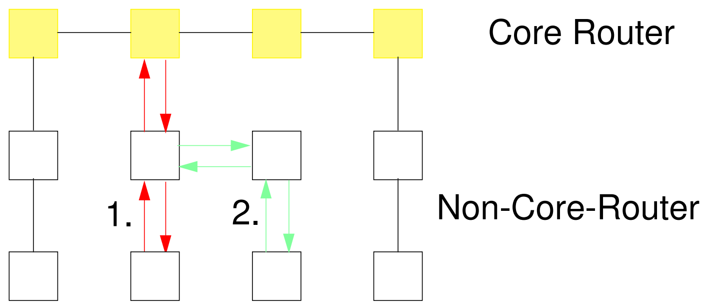
European Networking Center

Heidelberg

IP Multicast

Oktober 1996 - 21

Core-Based Tree (CBT)



- Host sendet IGMP Host Membership zu nächstem Router
- Router sendet Join_Request an den nächsten Core Router (Unicast)
- Core Router oder anderer Router des Multicast-Baums antworten mit Join_Ack
  -> Etablieren eines Zweiges des Multicast-Baums
- “Einklinken” in den Multicast-Baum bei weiteren Join_Request-Nachrichten (s. 2.)
- 1 gemeinsamer Baum für alle n Sender (statt n Bäume)

IBM

Dr. Torsten Braun

European Networking Center

Heidelberg

IP Multicast

Oktober 1996 - 22

Protocol Independent Multicast (PIM)

- Multicast-Routing-Protokolle berechnen und verteilen Multicast-spezifische Routing-Informationen.
- PIM basiert auf dem verfügbaren Unicast-Routing-Protokoll, d.h. es benutzt die Unicast-Routing-Information, um den Pfad zu einer Quelle zu bestimmen.
- PIM besitzt zwei Modi:
  - sparse Mode (PIM-SM) [sparse: “verstreut”]
  - dense mode (PIM-DM) [dense: “dicht”]
- Vorteile:
  - sehr einfaches Protokoll
  - kein Verteilen von Routing-Informationen
    -> Unabhängigkeit vom Unicast-Routing-Protokoll
- PIM-SM basiert auf CBT.
- PIM-DM basiert auf RPF.

Dr. Torsten Braun
European Networking Center
Heidelberg

PIM: Dense Mode (DM)

1) Router empfängt Multicast-Paket von S an G

2a) Falls Eingangs-Interface = Ausgangs-Interface für Unicast-Pakete an S: Paket wird auf allen anderen Interfaces weitergeleitet, sofern von diesen keine Pruning-Nachricht empfangen wurde

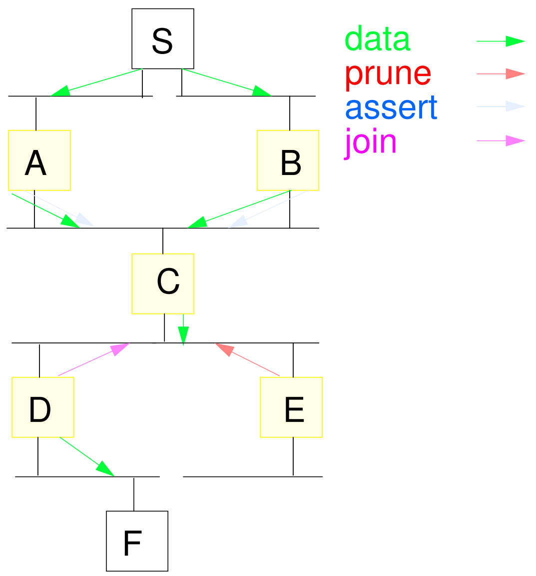
2b) sonst: Pruning-Nachricht prune(S,G) an Eingangs-Interface

- prune(S, G) wird auch gesendet, falls keine Gruppenmitglieder in einem Blatt-Link existieren
- graft(S,G) hebt Pruning wieder auf, Pruning-Informationen werden auch periodisch gelöscht.
- join(S,G): zum Aufheben von prune(S,G) durch Nachbar-Router am selben Link.
- assert(S,G): Anzeige der Distanz eines Routers zum Sender zur Kollisionserkennung

Unterschied zu DVMRP: zunächst werden Pakete über alle Interfaces weitergeleitet. Pruning reduziert den Multicast-Baum.

Dr. Torsten Braun
European Networking Center
Heidelberg

PIM-DM: Protokollablauf



assert => folgende Pakete werden nur über A nach C weitergeleitet

prune, join => C leitet Pakete nur an D weiter, nicht an E

IBM

Dr. Torsten Braun

European Networking Center

Heidelberg

PIM: Sparse Mode (SM)

- PIM-SM basiert auf dem CBT-Konzept.
- Zu jeder Gruppe gehören ein oder mehrere Rendezvous-Punkte (RP).
- Designated-Router (DR) stellen Verbindung zu RPs her Funktionen:
  - Empfangen
    - PIM Join/Prune-Nachricht an RP
    - RP überprüft periodisch die Erreichbarkeit der Router mit RP-Reachability-Nachricht
  - Senden
    - RP empfängt von Daten, die ggf. in PIM-Register-Nachricht eingekapselt sind
    - ggf. Auskapseln und Verteilen der Daten entlang des Multicast-Baums

IBM

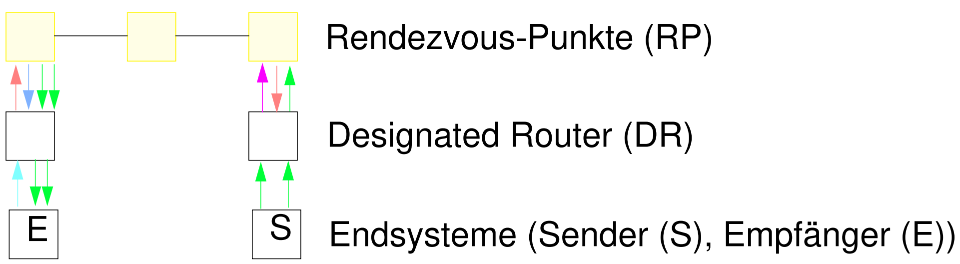
Dr. Torsten Braun

European Networking Center

Heidelberg

PIM-SM: Protokollablauf

- Beitreten eines Empfängers:
  - IGMP Membership Report und IGMP RP Report an DR
  - PIM-Join/Prune an RP
  - RP antwortet mit RP-Reachability
- Senden an Multicast-Gruppe:
  - Einkapseln des 1. Datenpakets in PIM-Register durch DR
  - RP sendet PIM-Join/Prune
  - DR leitet nachfolgende Daten ohne Einkapseln weiter
  - Verteilen der Daten durch RP



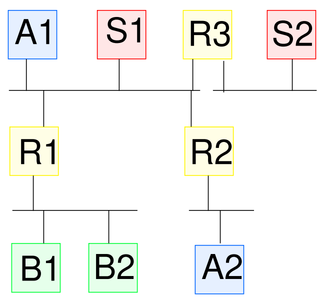
Dr. Torsten Braun
European Networking Center
Heidelberg

Multicast Open Shortest Path First (MOSPF)

Multicast-Erweiterung von OSPF: group-membership-LSAs (LSA: link state advertisement): [group, attached network], d.h. an welchen Links existieren Mitglieder welcher Gruppe.

Router gewinnen die Information aus den IGMP Membership Reports und verbreiten diese innerhalb einer Domäne.

Multicast-Pakete werden entsprechend des SPF-Algorithmus von Dijkstra weitergeleitet (Multicast-Baum f. jede Kombination von Sender/Gruppe)



R: Router
S: Sender
A: Mitglied der Gruppe A
B: Mitglied der Gruppe B

S2->B: R3 leitet Paket weiter
S1->B: R3 leitet Paket nicht weiter
S1->B: R1 leitet Paket weiter
S1->A: R1 leitet Paket nicht weiter

Dr. Torsten Braun
European Networking Center
Heidelberg

Multicast-Backbone (Mbone)

- Inbetriebnahme des ersten Multicast-Routers 1992
- erste Anwendung: Übertragung einer IETF Konferenz im Jahr 1992
- basiert auf IGMP (Hosts) und DVMRP (Multicast-Router)
- Aufbau eines Multicast-Overlay-Netzes
  -> Verbindung von Multicast-Routern über Tunnel-Technik, falls diese nicht an einen gemeinsamen Link angeschlossen sind.
- Mbone-Anwendungen:
  Video-, Audio-Konferenz-Werkzeuge, shared whiteboards

IBM

Dr. Torsten Braun

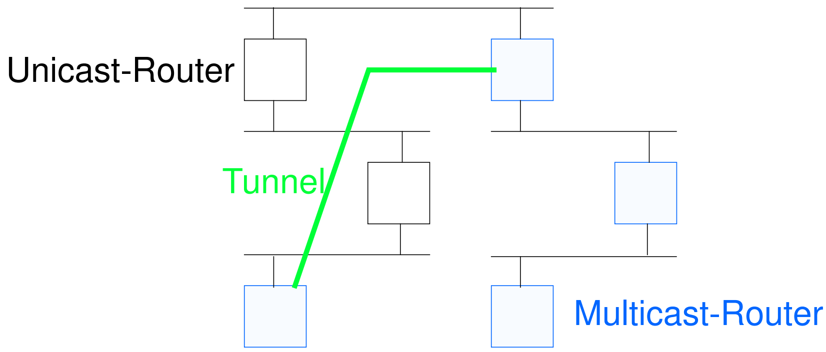
European Networking Center

Heidelberg

IP Multicast

Mbone-Topologie

- automatische Erkennung benachbarter Multicast-Router
- manuelle Konfiguration von Tunnels zwischen über Unicast-Router verbundene Multicast-Router
- Subnetze ohne Multicast-Router sind nicht ans Mbone angeschlossen.



- Tunneln: IP-Multicast-Paket wird am Tunnel-Anfang (IP-Quelladresse) in IP-Paket eingekapselt und am Tunnel-Ende (IP-Zieladresse) wieder ausgepackt.

IBM

Dr. Torsten Braun

European Networking Center

Heidelberg

IP Multicast

© IBM Deutschland 1996

Multicast-Anwendungen im Mbone

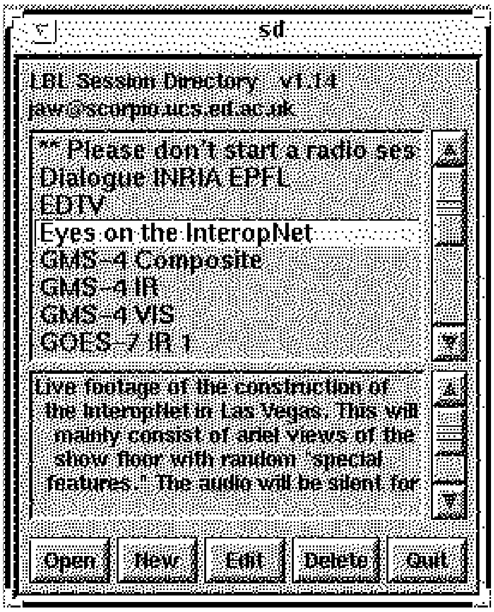
- Video-Konferenzen:
  - video conferencing tool (vic)
  - INRIA video conferencing system (ivs, incl. Audio)
  - network video (nv)
- Audio-Konferenzen:
  - visual audio tool (vat)
  - network audio conferencing/telephony tool (nevot)
  - robust audio tool (rat)
- Shared Editing:
  - whiteboard (wb)
  - Mbone text editor (nt)
- Session control:
  - session directory (sd, sdr)

Dr. Torsten Braun

European Networking Center

Heidelberg

Session Directory (sd)



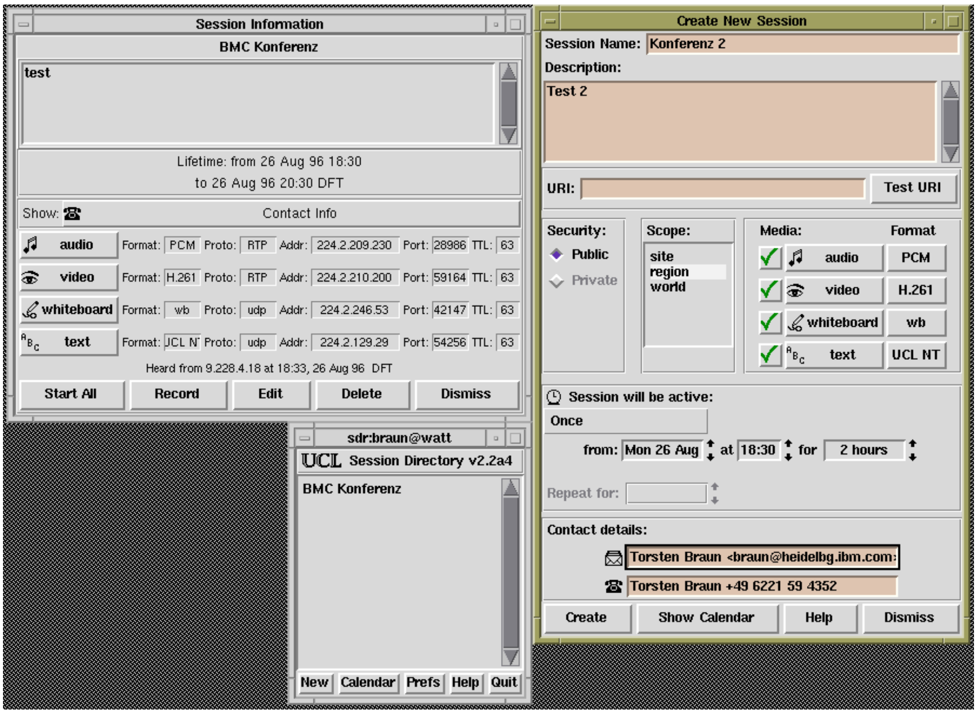
IP Multicast

Dr. Torsten Braun

European Networking Center

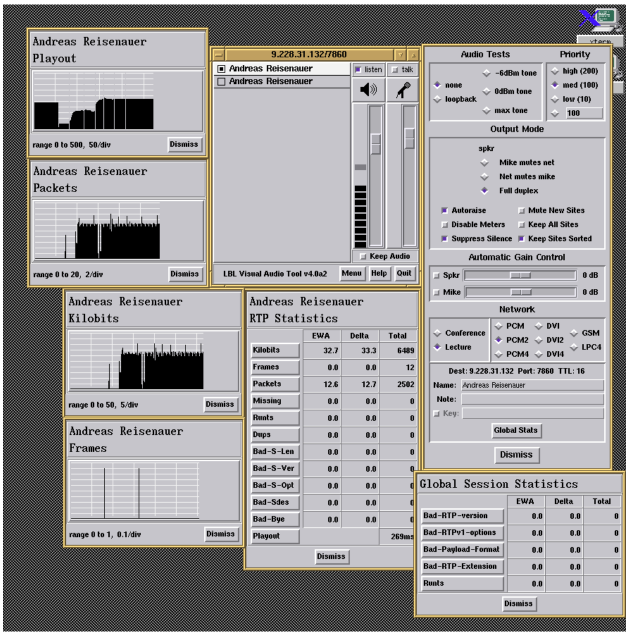
Heidelberg

Session Directory (sdr)

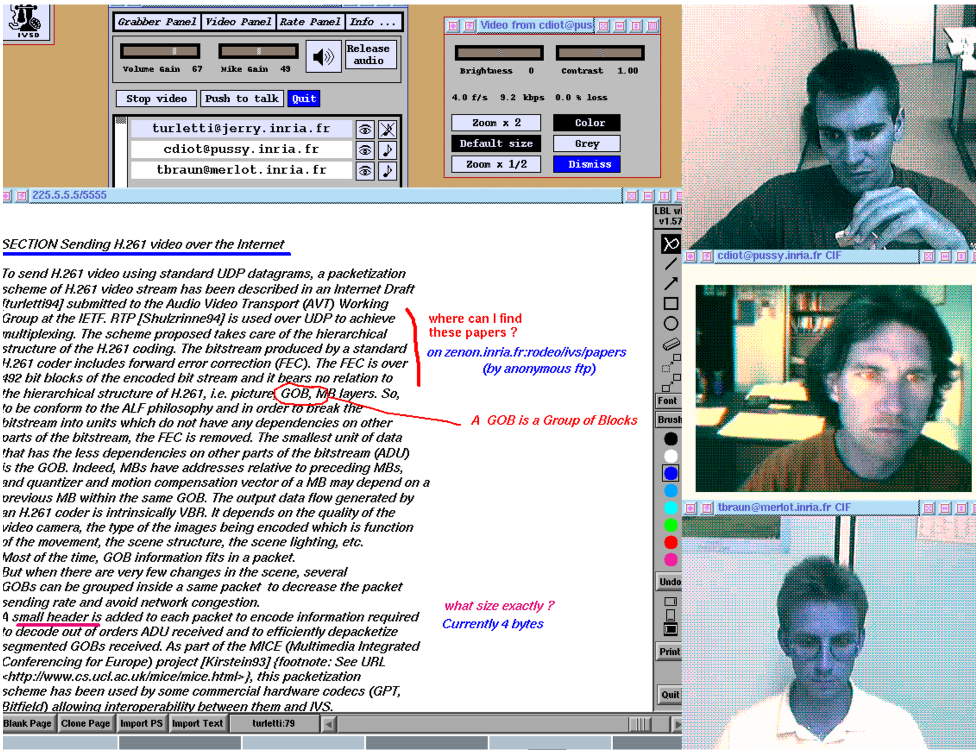


IP Multicast

Visual Audio Tool (vat)



Whiteboard und IVS



Voraussetzungen für Mbone-Anschluß

- IP-Multicast-Host-Erweiterungen (IGMP), in neuen Betriebssystemen (z.B. Windows 95, AIX 4) enthalten
- Multicast-Router im IP-Subnetz mit Tunnel an einen anderen Multicast-Router im Mbone
- Anwendungen (vic, vat, wb) existieren für alle wichtigen UNIX-Workstations und Windows 95
- Zunehmend auch kommerzielle Produkte über IP-Multicast

IBM
Dr. Torsten Braun
European Networking Center
Heidelberg

Literatur

- S. E. Deering: Multicast Routing in Internetworks and Extended LANs, ACM Sigcomm 1988, S. 55-64
- S.E. Deering: Host Extensions for IP Multicasting, RFC 1112
- D. Waitzmann, C. Partridge, S. Deering: Distance Vector Multicast Routing Protocol, RFC 1075
- C. Semeria, T. Maufer: Introduction to IP Multicast Routing, Internet Draft, ftp://nic.nordu.net/internet-drafts/draft-rfced-info-semeria-00.txt
- C. Huitema: Routing in the Internet, Prentice Hall 1995
- T. Braun, M. Zitterbart: Hochleistungskommunikation, Band 2: Transportdienste und -protokolle, Oldenbourg-Verlag 1996