

IP-Multicast / Mbone

Vorlesung

Rechnernetze und Internet – Fortgeschrittene Themen

13. Mai 2004

SS 2004 – Veranstaltungsnummer 260130

Guido Wessendorf

Zentrum für Informationsverarbeitung

Westfälische Wilhelms-Universität Münster

E-Mail: wessend@uni-muenster.de

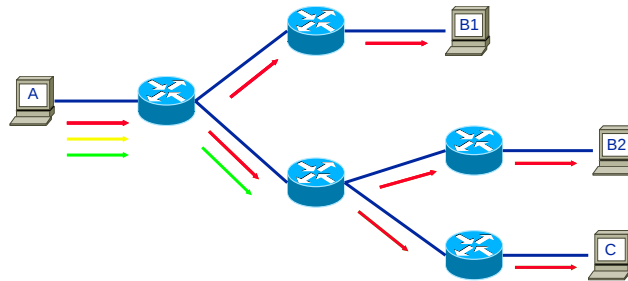


Themen

- Einführung
- IGMP
- Multicast Forwarding
- Mbone
- IP Multicast – Anwendungen

Einführung – IP Multicast

- 3 IP-Adress-Typen
 - Unicast (One-to-One)
 - Broadcast (One-to-All)
 - Multicast (One-to-Many)
- Gruppenkommunikation
- IP-Multicast wird (ist?) Schlüsseltechnologie im Internet !



IP Multicast-Adressen

- Multicast-IP-Paket nur durch Ziel-Adresse im Header erkennbar
- IP Multicast-Adresse = Gruppenadresse
- Class-D Adressen:



– 224.0.0.0 - 239.255.255.255 = 224.0.0.0/4

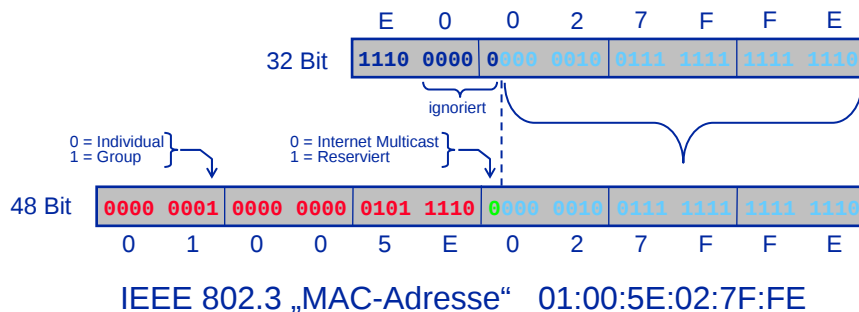
Gruppen-Adressen

- reservierte Adressen (IANA)
 - 224.0.0.0 - 224.0.0.255
 - » Beispiele
 - 224.0.0.1: alle Multicast-Systeme
 - 224.0.0.2: alle Router
 - 224.0.0.12: alle DHCP-Server/Relays
- limitierter Gültigkeitsbereich (Scope)
 - 239.0.0.0 - 239.255.255.255
 - » Beispiele
 - 239.192.0.0/14: Organization-Local-Scope
 - 239.255.0.0/16: Site-Local-Scope

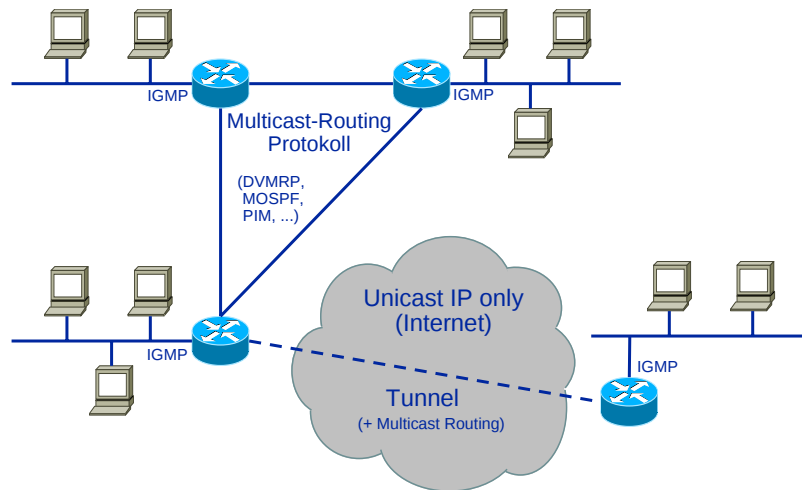
IP Multicast → MAC Adressen

- IANA reserviert MAC-Präfix 01:00:5E
 - Beispiel:

Class-D Adresse 224.2.127.254 (SAPv1)



Multicast Service-Modell



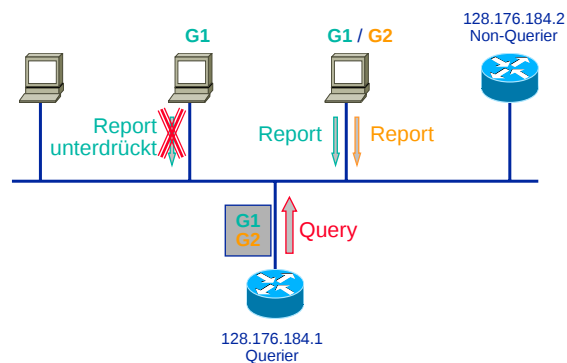
IGMP

- Internet Group Management Protocoll
 - IGMPv1 (RFC 1112)
 - IGMPv2 (RFC 2236, noch weitestgehend verbreitet)
 - IGMPv3 (RFC 3376, ersetzt RFC 2236)
- Hosts berichten jeweiligem Router über eigene Gruppenmitgliedschaft(en)
- Router sammeln Daten über Gruppenmitgliedschaften von direkt angebundenen Hosts

IGMPv1

- Host Membership Query
 - für jedes Subnetz sendet ausgewählter Router (Querier) Query-Messages an „All Systems Group“ (224.0.0.1) mit TTL=1
 - Query-Intervall zwischen 60 und 120 Sekunden
- Host Membership Report
 - Host antwortet für jede Gruppenmitgliedschaft mit Report-Message
 - nur ein Report pro Gruppe pro Subnetz:
 - » initialer Report-Delay
 - » erster Report unterdrückt Reports anderer Hosts
 - » Host schickt unaufgefordert Report beim Gruppeneintritt

IGMP-Query



1. Router sendet regelmäßig IGMP-Query
2. ein Host pro Gruppe pro LAN antwortet mit Report
3. Router verwaltet lokale Gruppen pro Interface

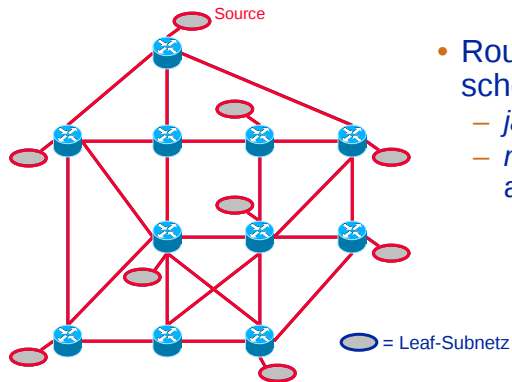
IGMPv2 und IGMPv3

- IGMPv2
 - Host sendet Leave-Message beim Verlassen einer Gruppe (vermeidet unnötigen Verkehr)
 - gruppenspezifisches Query möglich
 - Querier-Bestimmung standardisiert (Router mit kleinster IP-Adresse ist Querier)
- IGMPv3
 - Group-Source-Report-Message
 - » explizites Anfordern/Ablehnen von Sendern (in IGMPv1/v2 werden immer alle Sender akzeptiert)
 - » Verminderung unerwünschten Verkehrs
 - » Unterstützt „IGMP-Snooping“ durch Reports an 224.0.0.22 (s.u.)

Multicast-Forwarding

- IGMP nur für „Buchhaltung“ beim Router. Keine Regelung der Paketverteilung zwischen Routern oder Netzwerken
- effizientes, stabiles Multicast-Routing nicht trivial!
- Algorithmen für Multicast-Forwarding
 - Flooding
 - Spanning Tree (ST)
 - Reverse Path Broadcasting (RPB)
 - Truncated Reverse Path Broadcasting (TRPB)
 - Reverse Path Multicasting (RPM)
 - Core-Based Tree (CBT)

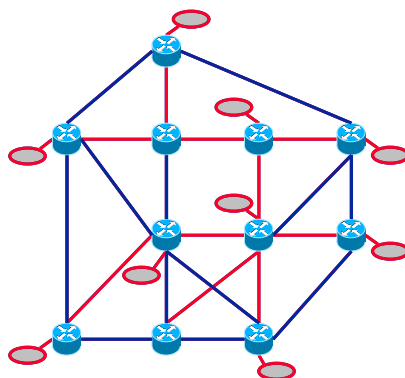
Flooding



- Router prüft, ob Paket schon mal gesehen
 - *ja*: Paket verwerfen
 - *nein*: Forward auf allen anderen Interfaces

- + simpel zu implementieren (keine Routingtabelle)
- viele duplizierte Pakete
- alle Wege werden immer benutzt
- Router müssen sich jedes Paket merken

Spanning Tree (ST)

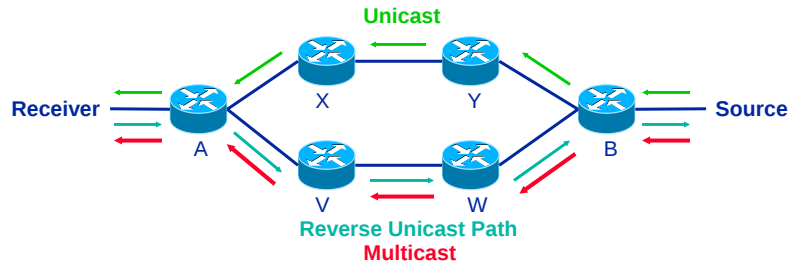


- nur *ein* aktiver Pfad verbindet zwei beliebige Router (keine Schleifen)
- Multicast-Pakete werden auf allen aktiven ST-Interfaces, ausgenommen dem Interface, von welchem die Pakete empfangen wurden, verschickt

- + Pakete werden nicht dupliziert
- Verkehr ggf. auf wenige Links konzentriert
- Wegewahl oft nicht optimal
- alle Subnetze werden mit einem ST adressiert

Reverse Path Broadcasting (RPB)

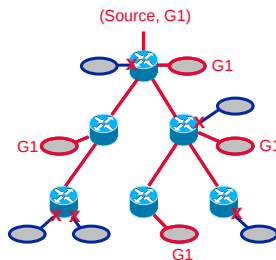
- für jedes aktive Paar (Source, Group) wird spezifischer „Spanning Tree“ (Shortest-Path-Tree, SPT) konstruiert:
 - Router leitet MC-Paket nur dann auf allen anderen Interfaces weiter, wenn es auf dem Interface empfangen wurde, welches der Router selber benutzen würde, um die Source zu erreichen.



- + Router braucht keine ST verwalten oder muss sich Pakete merken
- + effiziente Unicast-Wege für Multicast
- Forwarding in *alle* potentiellen Empfänger-Subnetze

Truncated Reverse Path Broadcasting

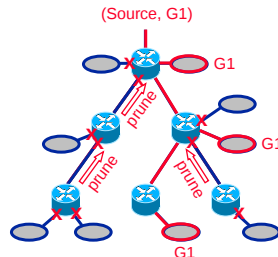
- mittels IGMP erkennen MC-Router lokale Gruppenmitgliedschaften in ihren Leaf-Subnetzen
- der SPT wird entsprechend abgeschnitten (truncated), wenn keine Empfänger vorliegen:



- + TRPB verbessert RPB: unnötiger Verkehr in Leaf-Netze wird unterdrückt
- beim Aufbau des SPT werden keine Gruppenmitgliedschaften berücksichtigt, d.h. alle MC-Router bekommen den Verkehr

Reverse Path Multicasting (RPM)

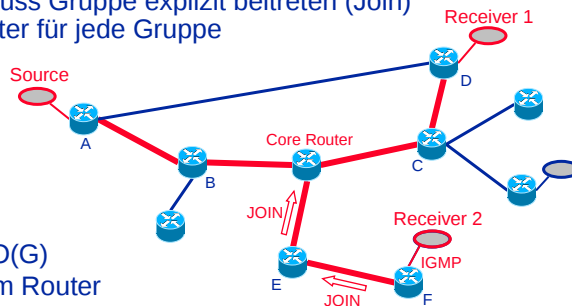
- Verbesserung von TRPB derart, dass der Auslieferungs-Baum „beschnitten“ wird (Pruning)
- MC-Verkehr nur noch
 - in Subnetzen mit Gruppenmitgliedern und
 - über Router/Subnetze entlang dem Weg zu Gruppenmitgliedern



- + MC-Verkehr nur dort, wo gewünscht bzw. notwendig
- jeder MC-Router muß $O(S,G)$ Zustände verwalten
- Gruppenmitgliedschaften/Netztopologien sind dynamischer Natur, also müssen regelmäßig Prunes kurzzeitig aufgehoben werden um ggf. neue Gruppenmitglieder bedienen zu können

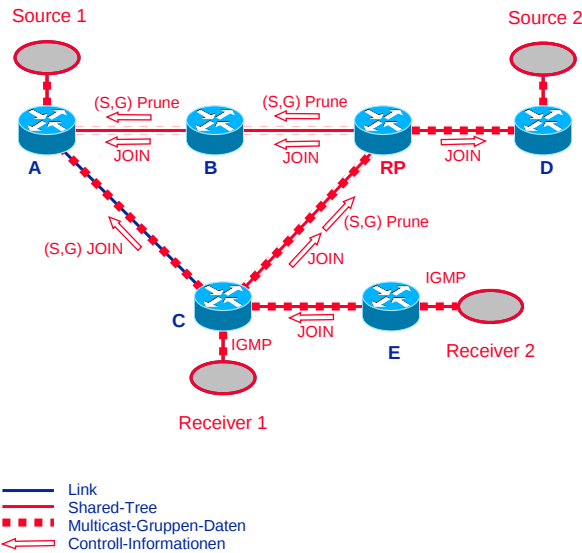
Core Based Tree (CBT)

- bisher: Source-Routed-Shortest-Path-Tree für jedes (S,G)-Paar.
- CBT: genau ein Auslieferungsbaum (Shared Tree) für jede Gruppe
 - von allen Sendern genutzt
 - Empfänger muss Gruppe explizit beitreten (Join)
 - ein Core-Router für jede Gruppe



- + skaliert besser:
statt $O(S,G)$ nur $O(G)$
Zustände in jedem Router
- + kein periodisches Flood&Prune
- Verkehrskonzentrationen im Core
- höhere Delays (suboptimale Routen)
- + CBTv3, PIM Sparse Mode (PIM-SM)

Beispiel: PIM sparse mode



1. Receiver 1 joins Gruppe G
2. RP erzeugt (*,G)-Status → C
3. Source 1 sendet Daten
4. A sendet Register zum RP. Register enthalten MC-Daten (encapsuliert). RP forwarded Daten.
5. RP sendet Join → A
6. A sendet Daten native, RP sendet Register-Stop
7. Verkehr (S1,G) an C überschreitet Schwellwert: Wechsel vom Shared-Tree zum Shortest-Path-Tree (S1,G)
8. C sendet Prunes entlang dem Shared-Tree zum Sender
9. Receiver 2 joins Gruppe G
10. Source 2 sendet Daten



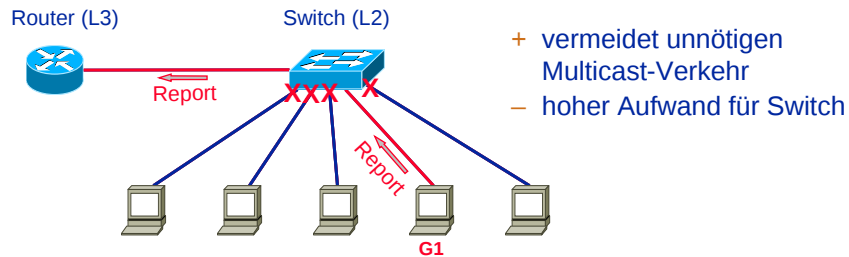
IP Multicast Routingprotokoll-Typen

- Dense Mode
 - nimmt „dichte“ Gruppenmitgliedschaft an
 - pruning dort, wo keine Teilnehmer
 - Beispiele:
 - » DVMRP („mrouted“)
 - » PIM-DM (Dense Mode)
 - » MOSPF
- Sparse Mode
 - Mitglieder über große Regionen dünn verteilt
 - Shortest-Path- oder Shared-Tree
 - explizites Join notwendig
 - Beispiele:
 - » CBT
 - » PIM-SM (Sparse Mode)
- Sparse-Dense Mode
 - „sinnvollste“ Methode wird verwendet ...
 - spezifischer Gruppen-Mode
 - falls Rendezvous-Point (RP) nicht bekannt: Dense Mode
 - Beispiel:
 - » PIM Sparse-Dense Mode



IGMP Snooping

- Level-2 Switches (Bridges) leiten IP-Multicast (Level 3) auf *allen* Interfaces weiter, da Pakete über Ethernet-Multicast transportiert werden
- IGMP Snooping: L2-Switch untersucht IGMP-Pakete (Membership Reports und Leave Messages) und ordnet Gruppenmitgliedschaften den jeweiligen Switch-Ports zu



- IGMPv3 unterstützt aktiv das Snooping von Layer-2 Switches durch Versenden von Gruppenreports an 224.0.0.22

IP Multicast – Historie

- grundlegende Arbeiten u.a. von Steve Deering Ende der 80er Jahre
- 1992
 - erste IETF-Experimente zur multimedialen Übertragung der IETF-Meetings
 - start des Mbone als virtuelles Multicast-Overlaynetz auf dem bestehenden Internet
 - etwa 40 Subnetze in 4 Ländern
- 1997
 - ca. 5000 (Sub-)Netze weltweit verteilt im Internet

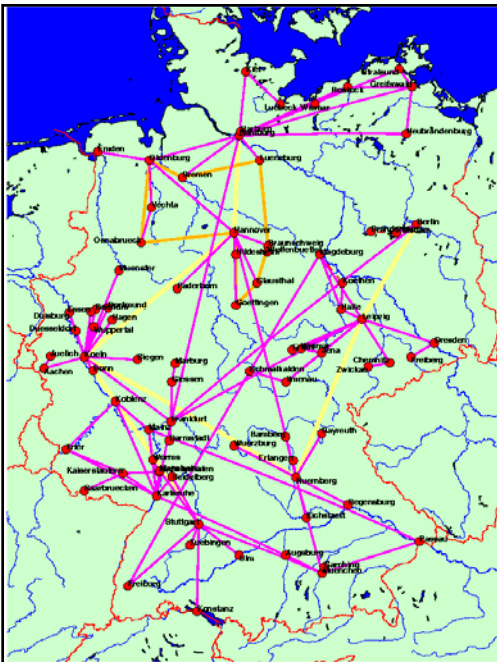


Abbildung: <http://www.rvs.uni-hannover.de/mbone/config/germany.html> (nicht mehr aktuell)

IP-Multicast/Mbone Grundlagen, 13. Mai 2004, Guido Wessendorf

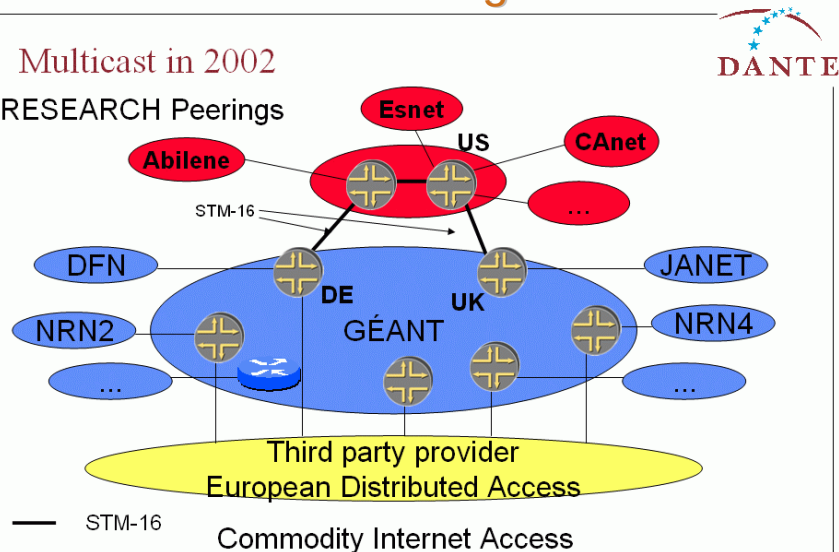
Mbone-DE DFN Multicast

- Standard Service im deutschen Wissenschaftsnetz (G-WiN)
- natives IP Multicast im Backbone (G-WiN) und bis zum Kundenrouter
- Protokolle
 - intern
 - » OSPF (Open Shortest Path First)
 - » PIM Sparse-Dense Mode
 - » Auto-RP
 - extern
 - » MBGP (Multicast BGP)
 - » PIM-SM
 - » MSDP (Multicast Source Discovery Protocol)
- ≈ 76 Teilnehmer

Internationale Anbindung

Multicast in 2002

RESEARCH Peerings



IP Multicast – Anwendungen



Frei verfügbare Mbone-Software

- weit verbreitet und angewendet
- für viele Betriebssysteme verfügbar
- Mbone-Tools, u.a.
 - sdr - Session Directory
 - vat - Audio (rat, fphone)
 - vic - Video
 - wb - Whiteboard (wbd)
 - nt - Text (nte)
 - MVoD - Conference Recording

sdr – Session Directory

IP-Multicast/Mbone Grundlagen, 13. Mai 2004, Guido Wessendorf

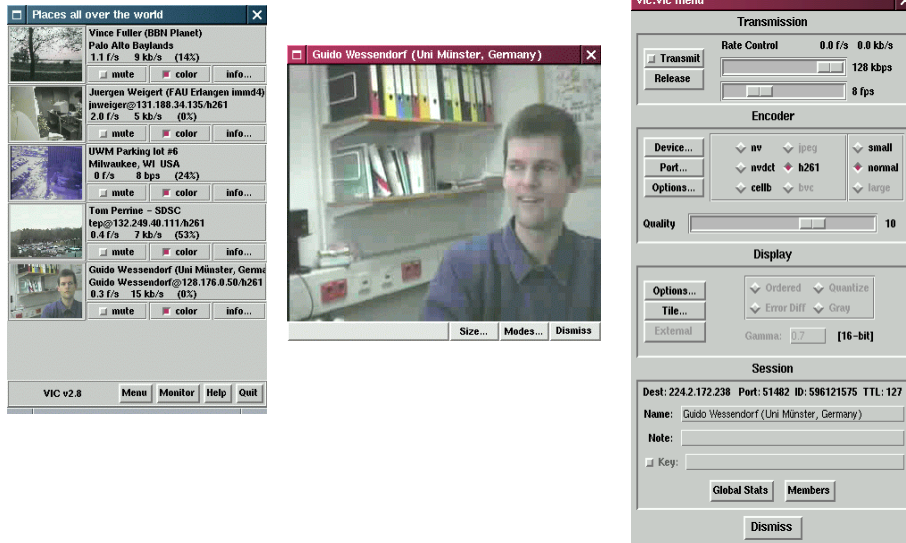
27

vat – Visual Audio Tool

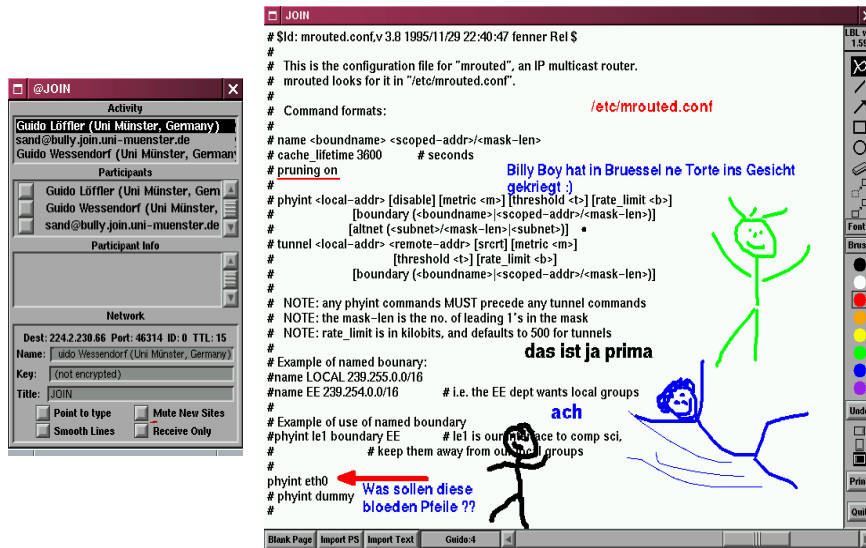
IP-Multicast/Mbone Grundlagen, 13. Mai 2004, Guido Wessendorf

28

vic – Video Conferencing Tool



wb – Shared Whiteboard



nt(e) – Network Text Editor

```
# $Id: routed.conf,v 3.8 1995/11/29 22:40:47 fenner Rel $
# This is the configuration file for "routed", an IP multicast router.
# routed looks for it in "/etc/routed.conf".

# Command formats:
# name <boundname> <scoped-addr>/<mask-len>
# cache_lifetime 3600 # seconds
# pruning on

# phyint <local-addr> [disable] [metric <m>] [threshold <t>] [rate_limit <b>]
# [boundary <boundname>|<scoped-addr>/<mask-len>]
# [altnet <subset>/<mask-len>|<subset>]
# tunnel <local-addr> <remote-addr> [srcrt] [metric <m>]
# [threshold <t>] [rate_limit <b>]
# [boundary <boundname>|<scoped-addr>/<mask-len>]

# NOTE: any phyint commands MUST precede any tunnel commands
# NOTE: the mask-len is the no. of leading 1's in the mask
# NOTE: rate_limit is in kilobits, and defaults to 500 for tunnels

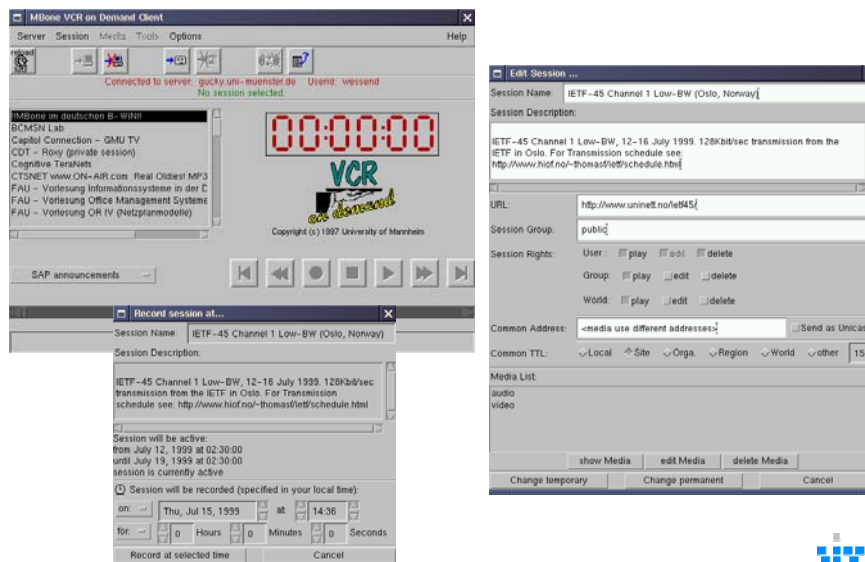
# Example of named boundary:
#name LOCAL 239.255.0.0/16
#name EE 239.254.0.0/16
# i.e. the EE dept wants local groups

# Example of use of named boundary
#phyint lei boundary EE
# lei is our interface to comp sci.
# keep them away from our local groups

#phyint eth0
# phyint dummy
# phyint eth0 disable (deaktivieren)

# Template tunnel for mcast install
#tunnel 128.4.0.77 128.4.0.8 metric 1 threshold 64 rate_limit 500 # <-- REPLACE
# boundary LOCAL
#tunnel 128.176.0.50 128.176.188.132 metric 1 threshold 1 rate_limit 0 (in LAN ok)
# Configuration des DWDM-Tunnels
# You might want to specify a boundary on your tunnel to the outside world,
# as above.
# NOTE: ONLY uncomment the following if you are running routed some!
```

MVoD – Conference Recording



Weitere Software

- Berkely MASH toolkit (Open Mash)
- Cisco IP/TV
- FreeAmp
- MacTV
- MIM
- MPlayer mit live.com patch
- Quicktime
- RealOne player
- VBrick Free Streaming player
- VLC Media Player
- WinAmp

(Kein Anspruch auf Vollständigkeit!)

Links

- mbone-de (DFN-NOC)
 - <http://www.mbone.de>
- Kompetenzzentrum für Videokonferenzdienste (VCC)
(DFN-Projekt an der TU-Dresden)
 - <http://vcc.urz.tu-dresden.de>
- University of Oregon videolab project
 - <http://videolab.uoregon.edu>