

Systeme II

4. Die Vermittlungsschicht

Christian Schindelhauer

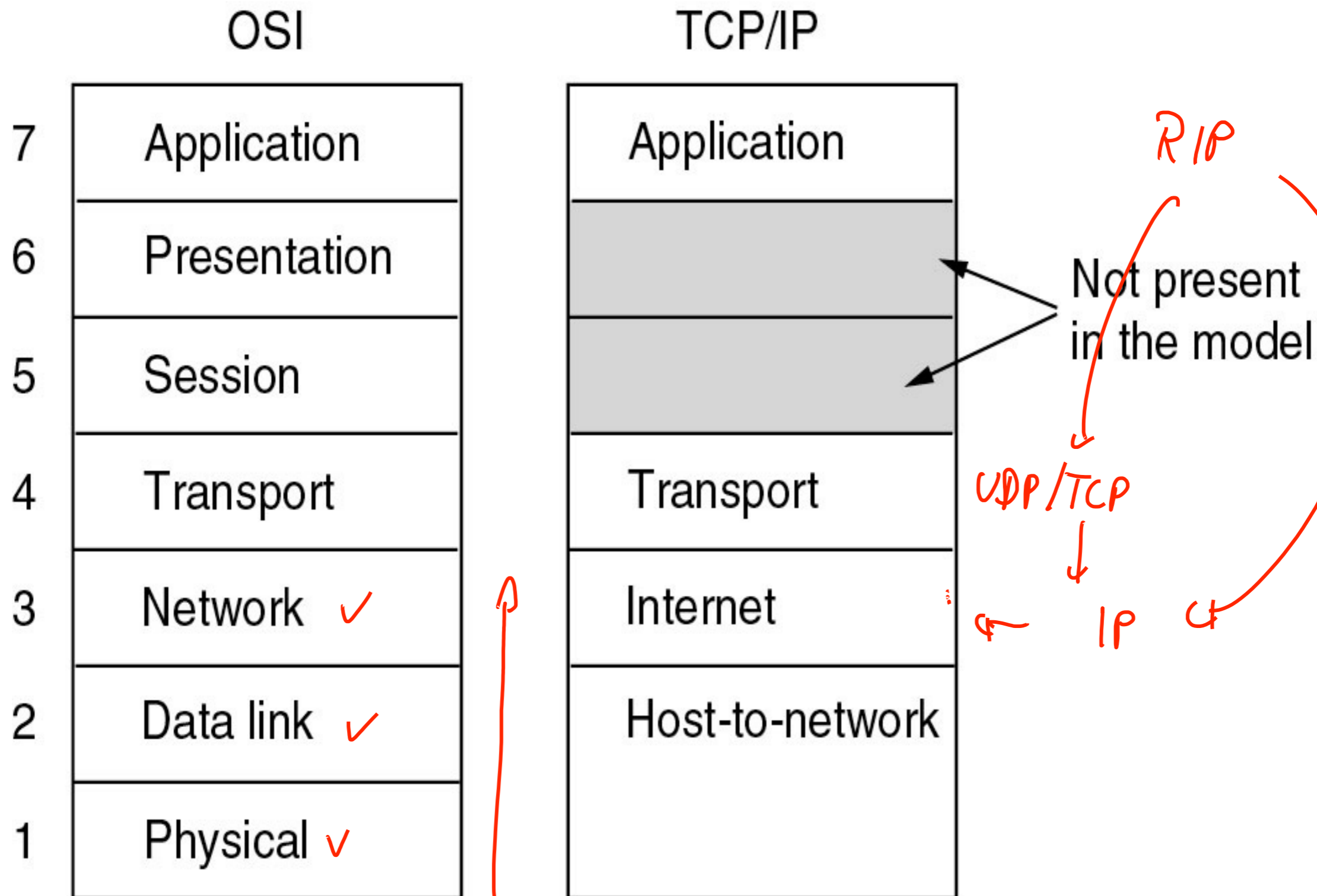
Technische Fakultät

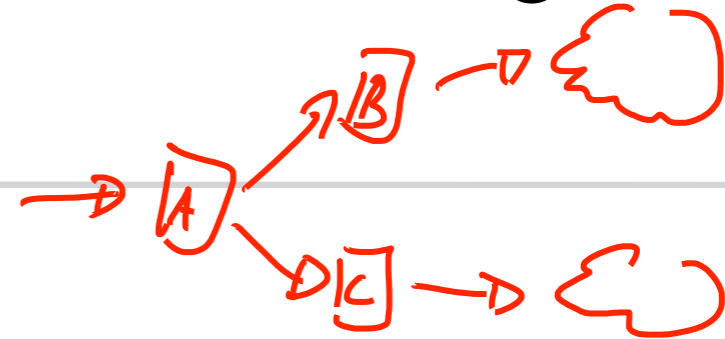
Rechnernetze und Telematik

Albert-Ludwigs-Universität Freiburg

Version 07.06.2016

OSI versus TCP/IP





■ IP-Routing-Tabelle

- enthält für Ziel (Destination) die Adresse des nächsten Rechners (Gateway)
- Destination kann einen Rechner oder ganze Sub-nets beschreiben
- Zusätzlich wird ein Default-Gateway angegeben

■ Packet Forwarding

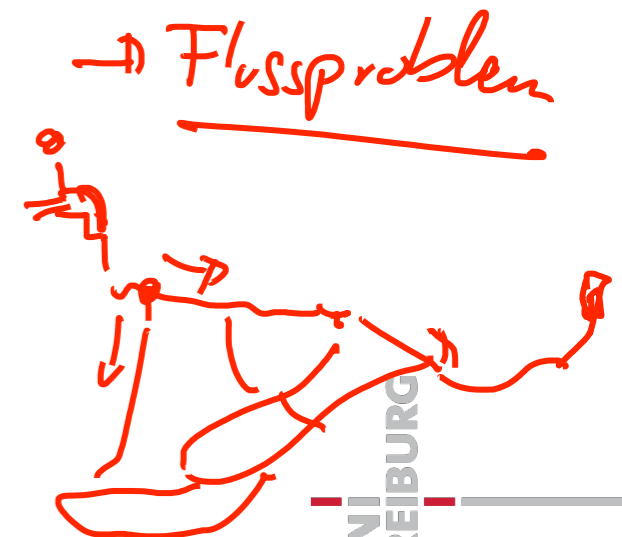
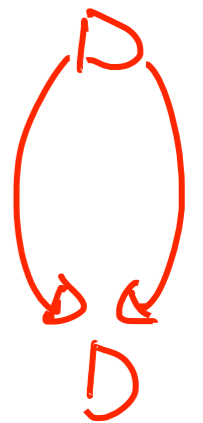
- früher Packet Routing genannt
- IP-Paket (datagram) enthält Start-IP-Adresse und Ziel-IP-Adresse

→ • Ist Ziel-IP-Adresse = eigene Rechneradresse dann Nachricht ausgeliefert

→ • Ist Ziel-IP-Adresse in Routing-Tabelle dann leite Paket zum angegebenen Gateway

→ • Ist Ziel-IP-Subnetz in Routing-Tabelle dann leite Paket zum angegebenen Gateway

→ • Ansonsten leite zum Default-Gateway



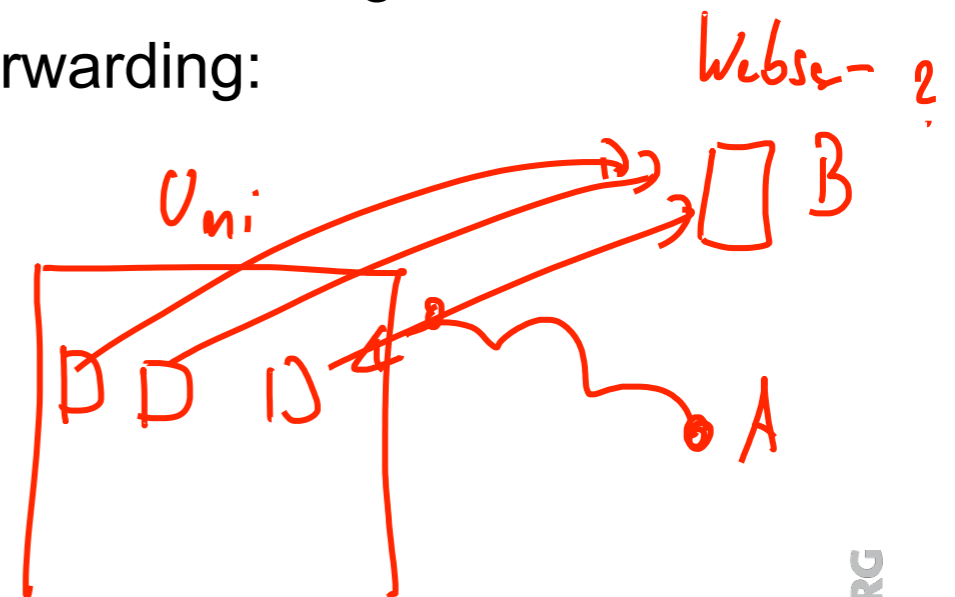
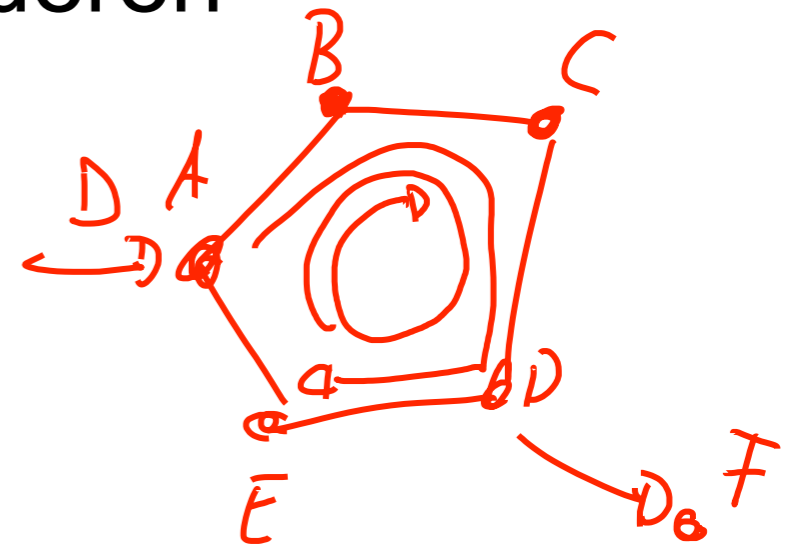
- IP-Paket (datagram) enthält unter anderen

- TTL (Time-to-Live): Anzahl der Hops
- Start-IP-Adresse
- Ziel-IP-Adresse

- **Behandlung eines Pakets**

- Verringere TTL (Time to Live) um 1
- ↳ Falls TTL $\neq 0$ dann Packet-Forwarding aufgrund der Routing-Tabelle
- ↳ Falls TTL = 0 oder bei Problemen in Packet-Forwarding:
 - Lösche Paket
 - Falls Paket ist kein ICMP-Paket dann
 - Sende ICMP-Paket mit
 - Start= aktuelle IP-Adresse und
 - Ziel = alte Start-IP-Adresse

traceroute tracert



- Forwarding:
 - Weiterleiten von Paketen
- Routing:
 - Erstellen Routen, d.h.
 - Erstellen der Routing-Tabelle
- Statisches Routing
 - Tabelle wird manuell erstellt
 - sinnvoll für kleine und stabile LANs
- Dynamisches Routing
 - Tabellen werden durch Routing-Algorithmus erstellt
 - Zentraler Algorithmus, z.B. Link State
 - Einer/jeder kennt alle Information, muss diese erfahren
 - Dezentraler Algorithmus, z.B. Distance Vector
 - arbeitet lokal in jedem Router
 - verbreitet lokale Information im Netzwerk

Distance Vector Routing Protocol

Edsger Wybe / Dijkstra

B	A	C	
A	2	-	A
C	-	1	C
E	11	-	

Distance Table Datenstruktur

- Jeder Knoten besitzt eine
 - Zeile für jedes mögliches Ziel
 - Spalte für jeden direkten Nachbarn

Verteilter Algorithmus

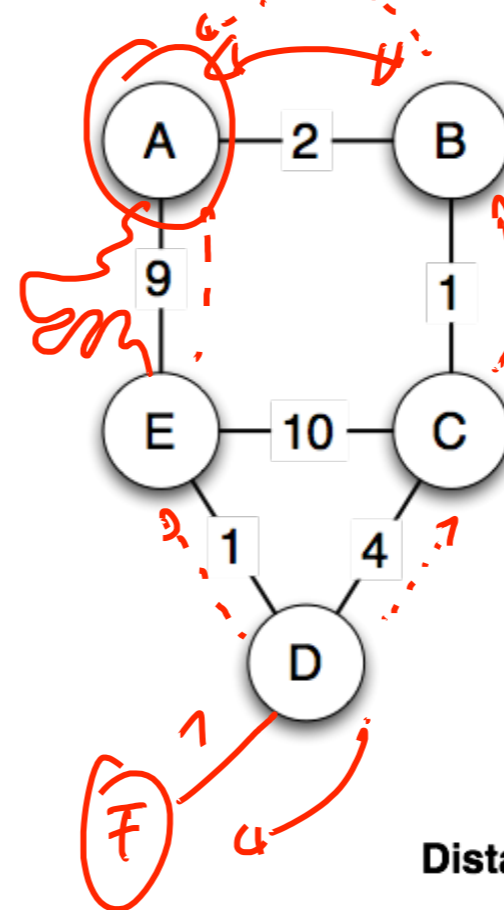
- Jeder Knoten kommuniziert nur mit seinem Nachbarn

Asynchroner Betrieb

- Knoten müssen nicht Informationen austauschen in einer Runde

Selbst Terminierend

- läuft bis die Knoten keine Informationen mehr austauschen



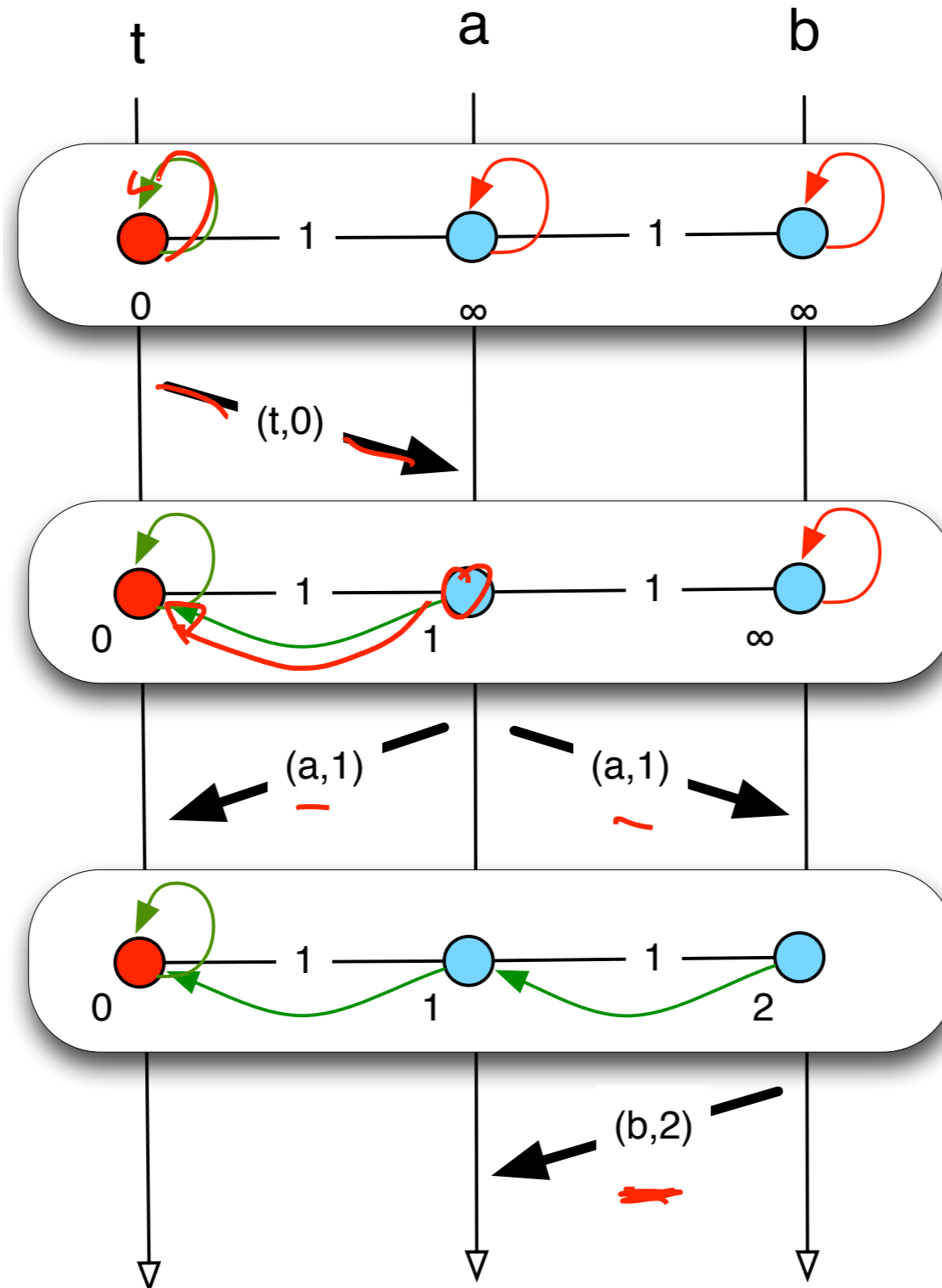
Distance Table für A

von A	über		Routing Tabellen Eintrag	
	B	E		
nach B	2	15	B	2
C	3	14	B	3
D	7	10	B	7
E	8	9	E	8

Distance Table für C

von C	über			Routing Tabellen Eintrag
	B	D	E	
nach A	3	11	18	B
B	1	9	16	B
D	6	4	11	D
E	7	5	10	D

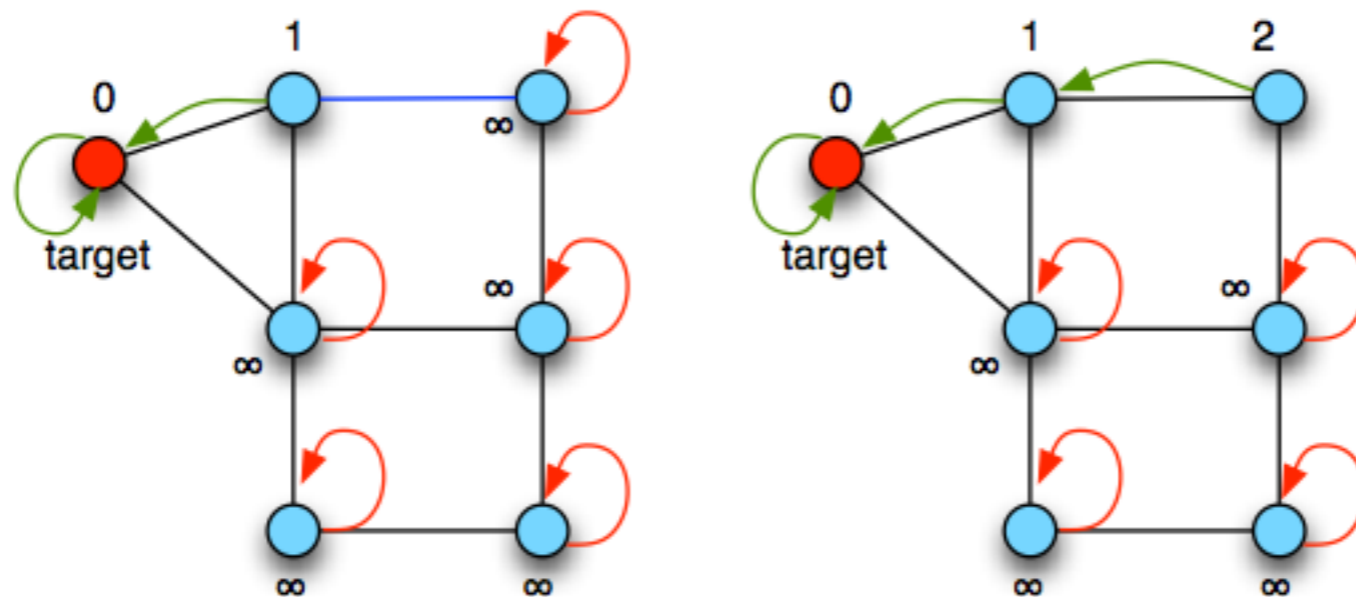
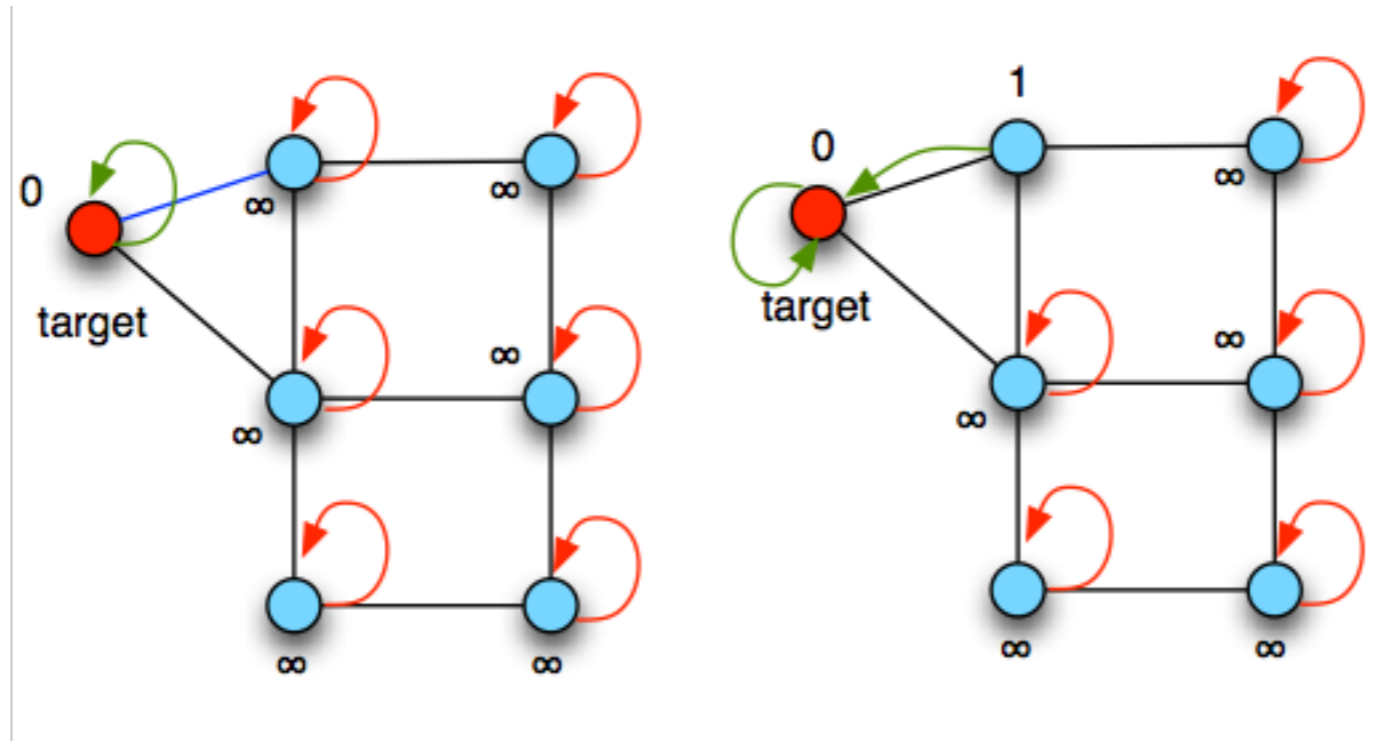
Beispiel für Distance-Vector für Ziel t



über

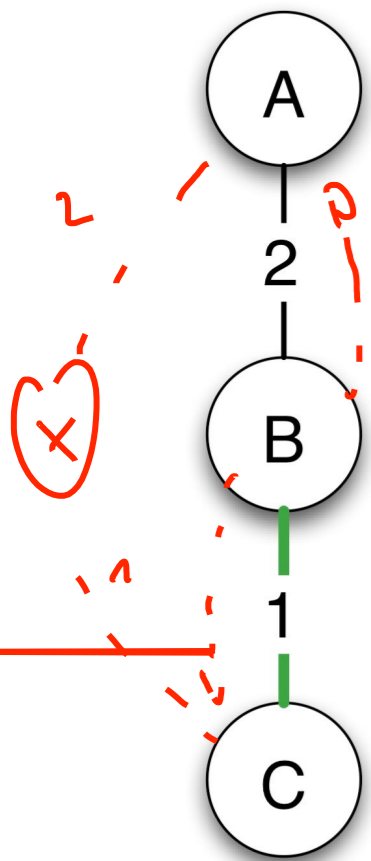
a	t	b
t	1	3

Distance-Vector für ein Ziel



Das “Count to Infinity” - Problem

- Gute Nachrichten verbreiten sich schnell
 - Neue Verbindung wird schnell veröffentlicht



Distance Table für A

		über		Routing Tabellen Eintrag
von A		B		
nach	B	2		B
	C	-		B

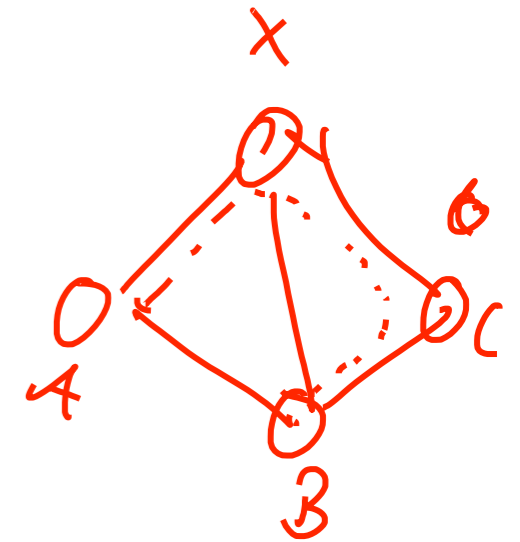
Distance Table für B

		über		Routing Tabellen Eintrag
von B		A	C	
nach	A	2	-	A
	C	-	1	C

nach kurzer Zeit

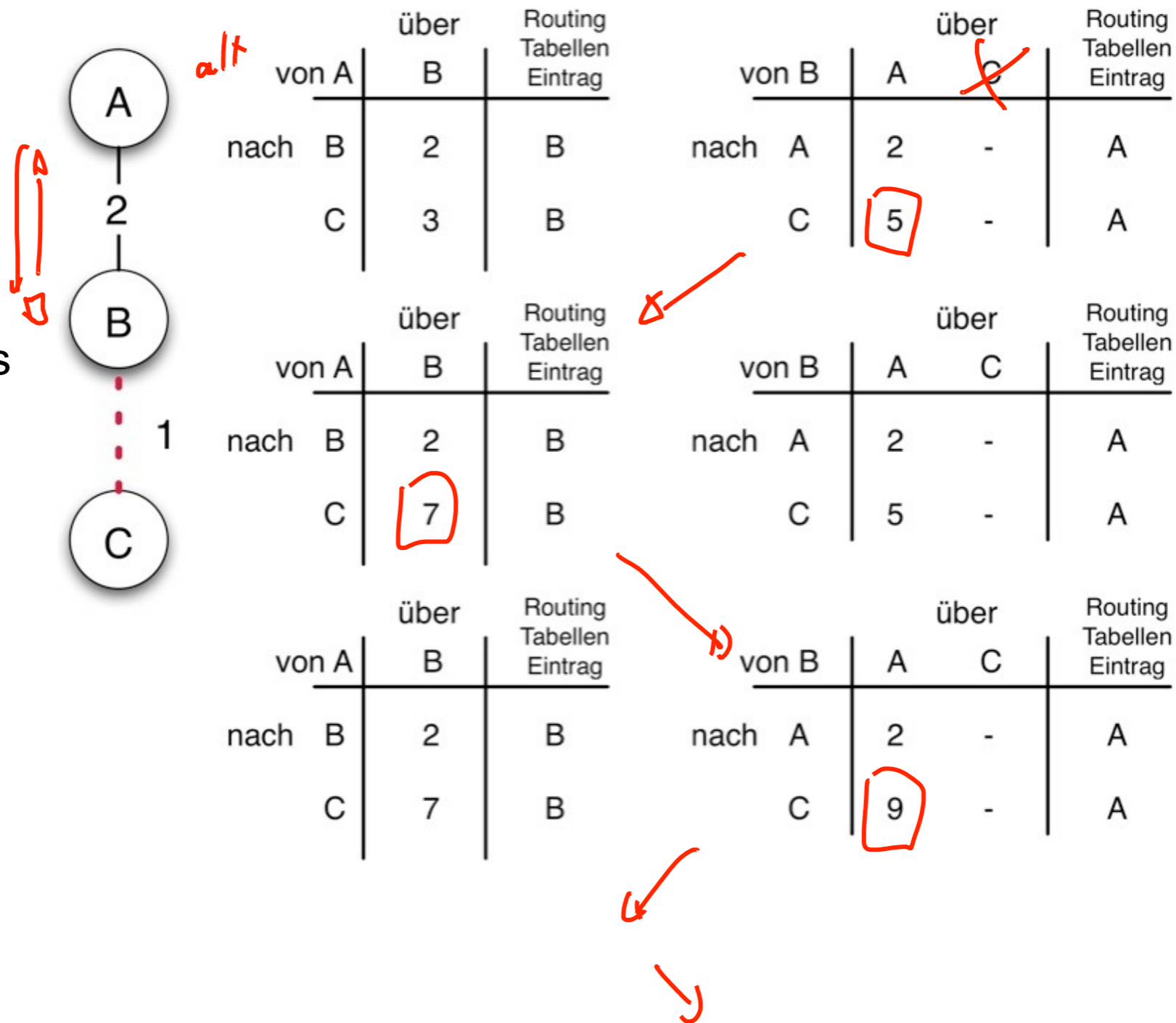
		über		Routing Tabellen Eintrag
von A		B		
nach	B	2		B
	C	3		B

		über		Routing Tabellen Eintrag
von B		A	C	
nach	A	2	4	A
	C	5	1	C

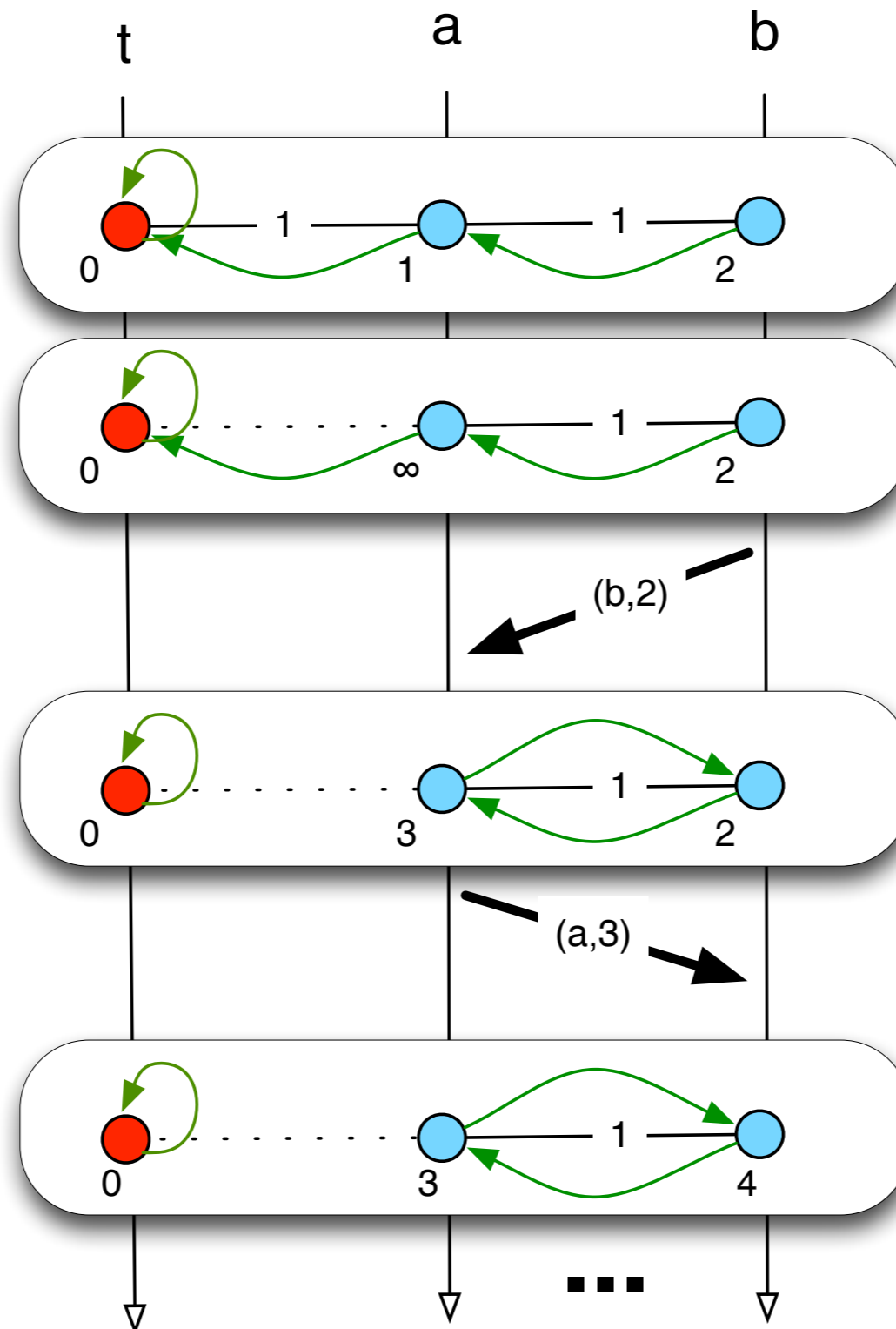


Das “Count to Infinity” - Problem

- Schlechte Nachrichten verbreiten sich langsam
 - Verbindung fällt aus
 - Nachbarn erhöhen wechselseitig ihre Entfernung
 - “Count to Infinity”-Problem



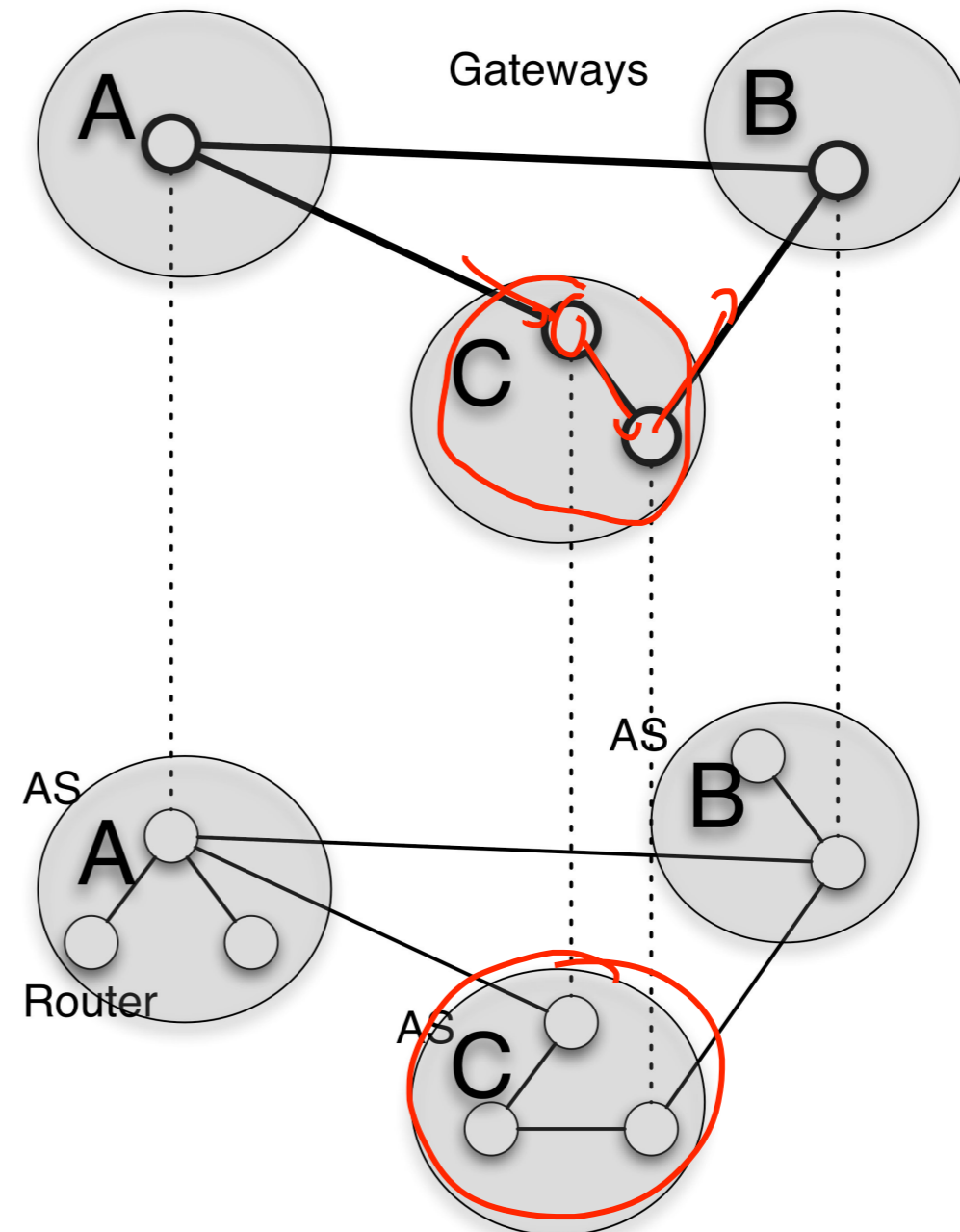
Das "Count to Infinity" - Problem für Ziel t



- Link State Router
 - tauschen Information mittels Link State Packets (LSP) aus
 - Jeder verwendet einen eigenen Kürzeste-Wege-Algorithmus zu Anpassung der Routing-Tabelle
- LSP enthält
 - ID des LSP erzeugenden Knotens
 - Kosten dieses Knotens zu jedem direkten Nachbarn
 - Sequenznr. (SEQNO)
 - TTL-Feld für dieses Feld (time to live)
- Verlässliches Fluten (Reliable Flooding)
 - Die aktuellen LSP jedes Knoten werden gespeichert
 - Weiterleitung der LSP zu allen Nachbarn
 - bis auf den Knoten der diese ausgeliefert hat
 - Periodisches Erzeugen neuer LSPs
 - mit steigender SEQNOs
 - Verringern der TTL bei jedem Weiterleiten

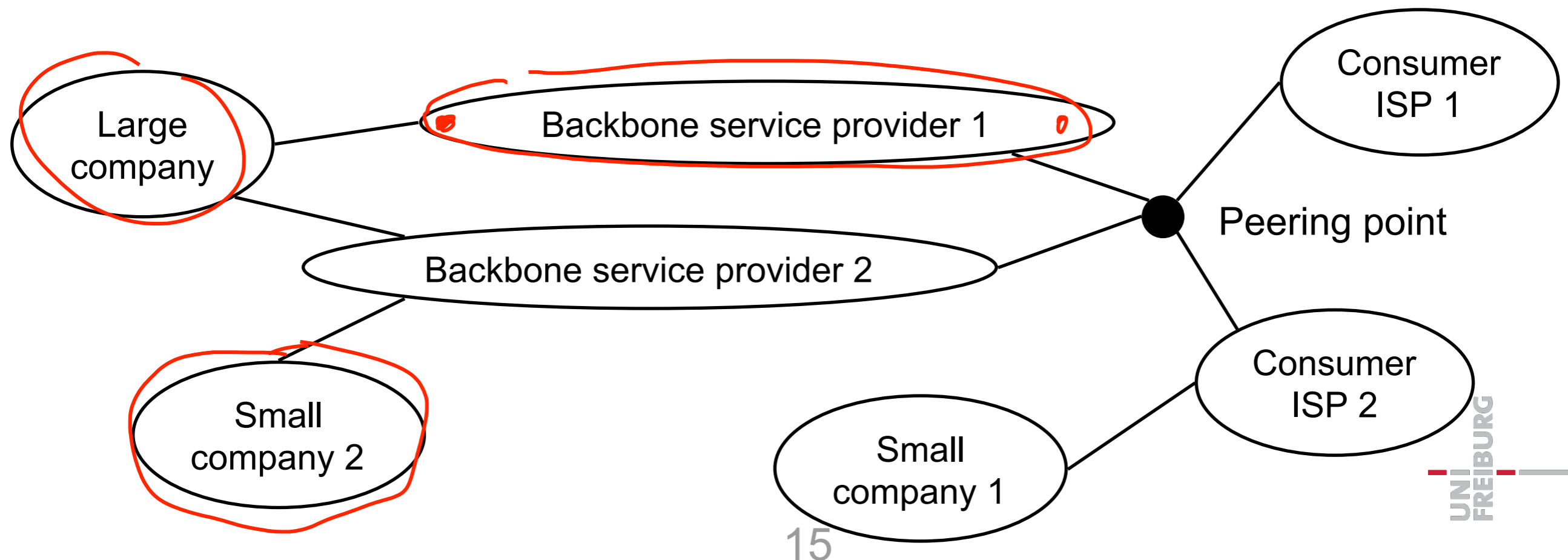
- Link State Routing
 - benötigt $O(g \cdot n)$ Einträge für n Router mit maximalen Grad g
 - Jeder Knoten muss an jeden anderen seine Informationen senden
- Distance Vector
 - benötigt $O(g \cdot n)$ Einträge
 - kann Schleifen einrichten
 - Konvergenzzeit steigt mit Netzwerkgröße
- Im Internet gibt es mehr als 10^7 Router
 - damit sind diese so genannten flachen Verfahren nicht einsetzbar
- Lösung:
 - Hierarchisches Routing

- **Autonomous System (AS)**
 - liefert ein zwei Schichten-Modell des Routing im Internet
 - Beispiele für AS:
 - uni-freiburg.de
- **Intra-AS-Routing (Interior Gateway Protocol)**
 - ist Routing innerhalb der AS
 - z.B. RIP, OSPF, IGRP, ...
- **Inter-AS-Routing (Exterior Gateway Protocol)**
 - Übergabepunkte sind Gateways
 - ist vollkommen dezentrales Routing
 - Jeder kann seine Optimierungskriterien vorgeben
 - z.B. EGP (früher), BGP

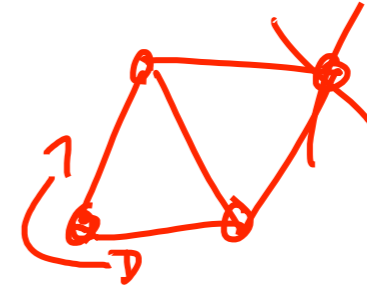


Typen autonomer Systeme

- ✗ Stub-AS
 - Nur eine Verbindung zu anderen AS
- Multihomed AS
 - Verbindungen zu anderen ASen
 - weigert sich aber Verkehr für andere zu befördern
- Transit AS
 - Mehrere Verbindungen
 - Leitet fremde Nachrichten durch (z.B. ISP)



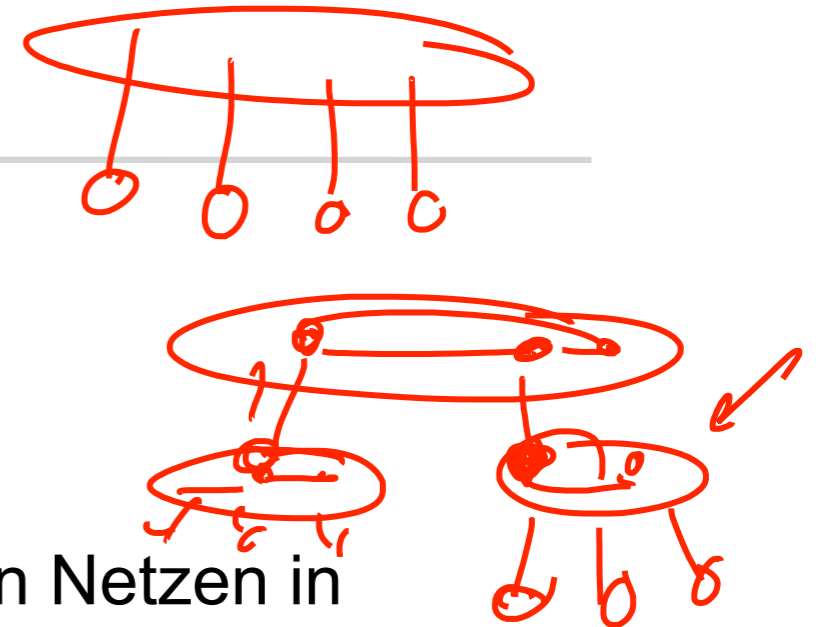
- Distance Vector Algorithmus
 - Distanzmetrik = Hop-Anzahl
- Distanzvektoren
 - werden alle 30s durch Response-Nachricht (advertisement) ausgetauscht
- Für jedes Advertisement
 - Für bis zu 25 Zielnetze werden Routen veröffentlicht per UDP
- Falls kein Advertisement nach 180s empfangen wurde
 - Routen über Nachbarn werden für ungültig erklärt
 - Neue Advertisements werden zu den Nachbarn geschickt
 - Diese antworten auch mit neuen Advertisements
 - falls die Tabellen sich ändern
 - Rückverbindungen werden unterdrückt um Ping-Pong-Schleifen zu verhindern (poison reverse) gegen Count-to-Infinity-Problem
 - Unendliche Distanz = 16 Hops



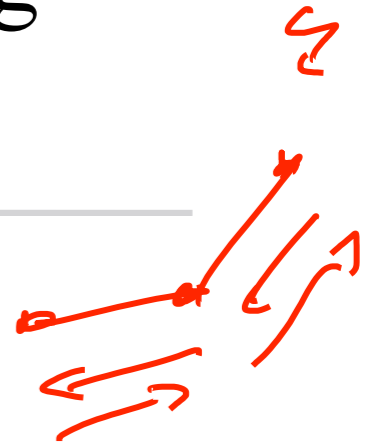
Intra-AS OSPF (Open Shortest Path First)

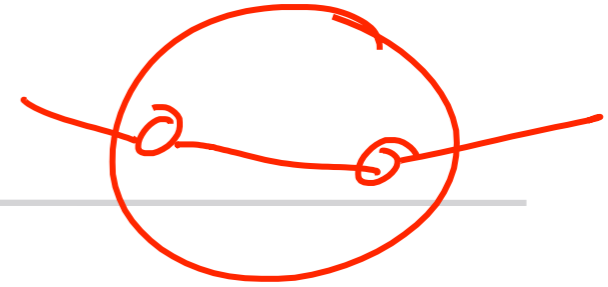
- “open” = öffentlich verfügbar
- Link-State-Algorithmus
 - LS Paket-Verbreitung
 - Topologie wird in jedem Knoten abgebildet
 - Routenberechnung mit Dijkstras Algorithmus
- OSPF-Advertisment
 - per TCP, erhöht Sicherheit (security)
 - werden in die gesamte AS geflutet
 - Mehrere Wege gleicher Kosten möglich

- Für große Netzwerke zwei Ebenen:
 - Lokales Gebiet und Rückgrat (backbone)
 - Lokal: Link-state advertisement
 - Jeder Knoten berechnet nur in Richtung zu den Netzen in anderen lokalen Gebieten
- Local Area Border Router:
 - Fassen die Distanzen in das eigene lokale Gebiet zusammen
 - Bieten diese den anderen Area Border Routern an (per Advertisement)
- Backbone Routers
 - verwenden OSPF beschränkt auf das Rückgrat (backbone)
- Boundary Routers:
 - verbinden zu anderen AS



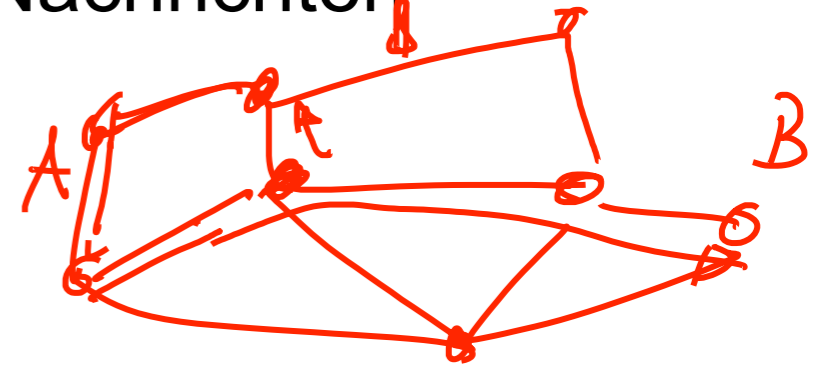
- CISCO-Protokoll, Nachfolger von RIP (1980er)
- Distance-Vector-Protokoll, wie RIP
 - Hold Down
 - weggefallene Verbindungen werden mit Entfernung „unendlich“ angeboten (100 = unendlich)
 - Split Horizon
 - Advertisements werden nicht auf dem angebotenen Pfad weitergegeben
 - Poison Reverse
 - weggefallene Verbindungen werden sofort mit Entfernung „unendlich“ allen Nachbarn angeboten
- Verschiedene Kostenmetriken
 - Delay, Bandwidth, Reliability, Load etc.
- Verwendet TCP für den Austausch von Routing Updates



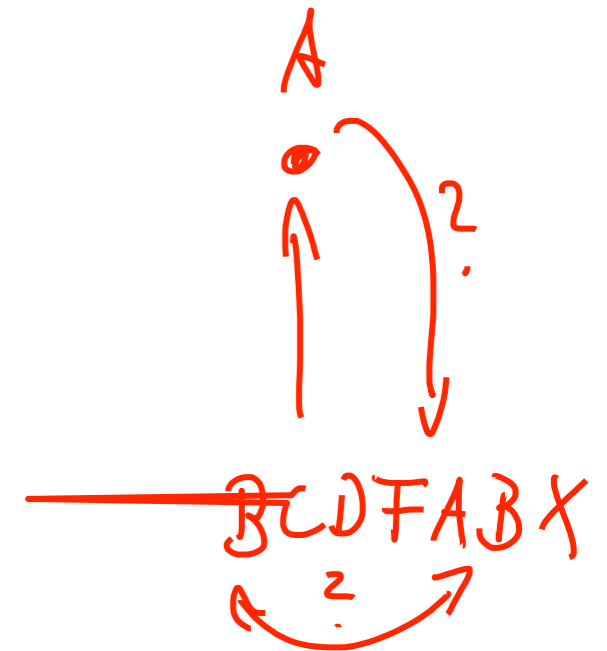
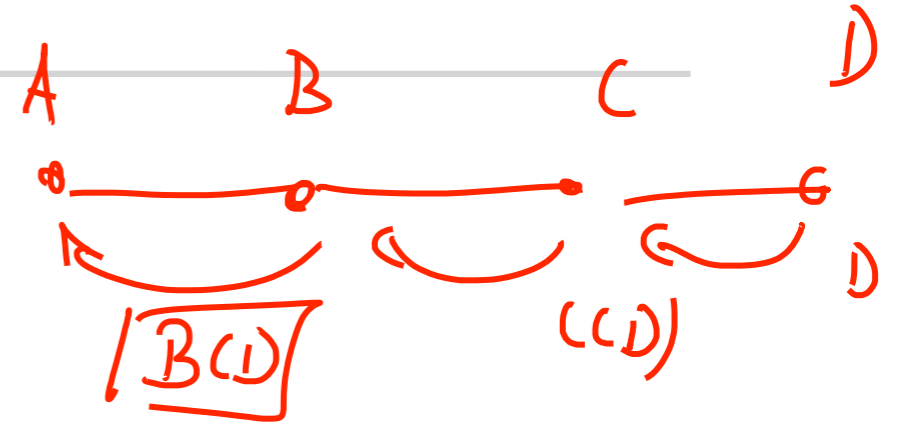


Inter-AS-Routing ist schwierig...

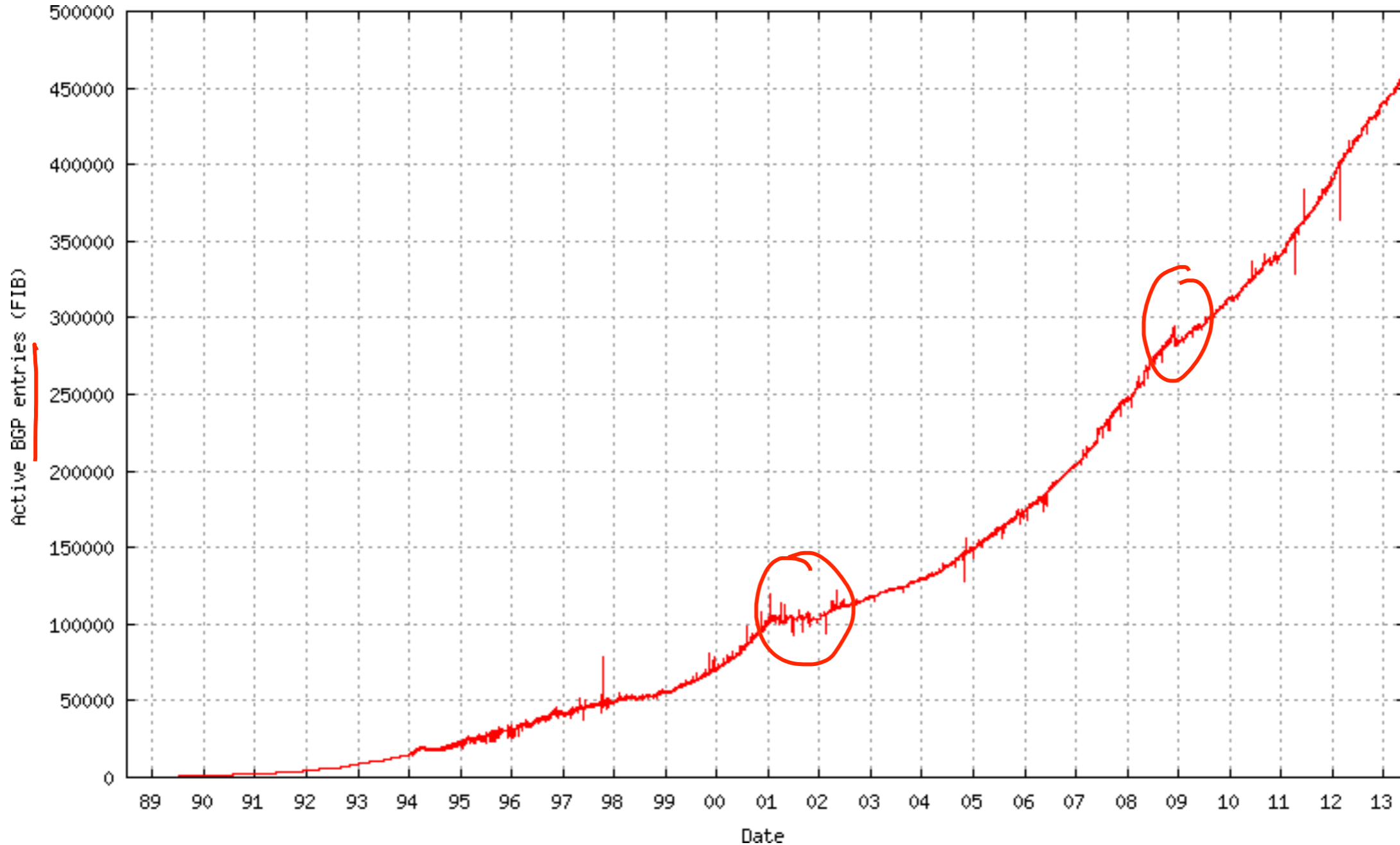
- Organisationen können Durchleitung von Nachrichten verweigern
- Politische Anforderungen
 - Weiterleitung durch andere Länder?
- Routing-Metriken der verschiedenen autonomen Systeme sind oftmals unvergleichbar
 - Wegeoptimierung unmöglich!
 - Inter-AS-Routing versucht wenigstens Erreichbarkeit der Knoten zu ermöglichen
- Größe: momentan müssen Inter-Domain-Router mehr als 300.000 Einträge verwalten



- Ist faktisch der Standard
- Path-Vector-Protocol
 - ähnlich wie Distance Vector Protocol
 - es werden aber ganze Pfade zum Ziel gespeichert
 - jeder Border Gateway teilt all seinen Nachbarn (peers) den gesamten Pfad (Folge von ASen) zum Ziel mit (advertisement) (per TCP)
- Falls Gateway X den Pfad zum Peer-Gateway W sendet
 - dann kann W den Pfad wählen oder auch nicht
 - Optimierungskriterien:
 - Kosten, Politik, etc.
 - Falls W den Pfad von X wählt, dann publiziert er
 - $\text{Path}(W,Z) = (W, \text{Path}(X,Z))$
- Anmerkung
 - X kann den eingehenden Verkehr kontrollieren durch Senden von Advertisements
 - Sehr kompliziertes Protokoll



BGP-Routing Tabellengröße 1994-2013



Broadcast & Multicast

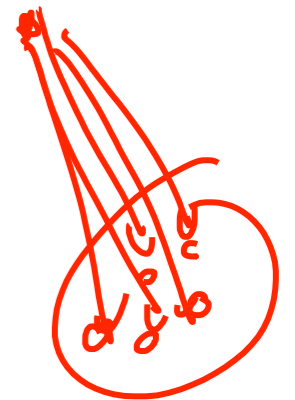


■ Broadcast routing →

- Ein Paket soll (in Kopie) an alle ausgeliefert werden
- Lösungen:
 - Fluten des Netzwerks
 - Besser: Konstruktion eines minimalen Spannbaums

■ Multicast routing

- Ein Paket soll an eine gegebene Teilmenge der Knoten ausgeliefert werden (in Kopie)
- Lösung:
 - Optimal: Steiner Baum Problem (bis heute nicht lösbar)
 - Andere (nicht-optimale) Baum-konstruktionen



Uunicast ~~point-to-point~~ | Any/all
Point-to-point | (subset)

- Motivation

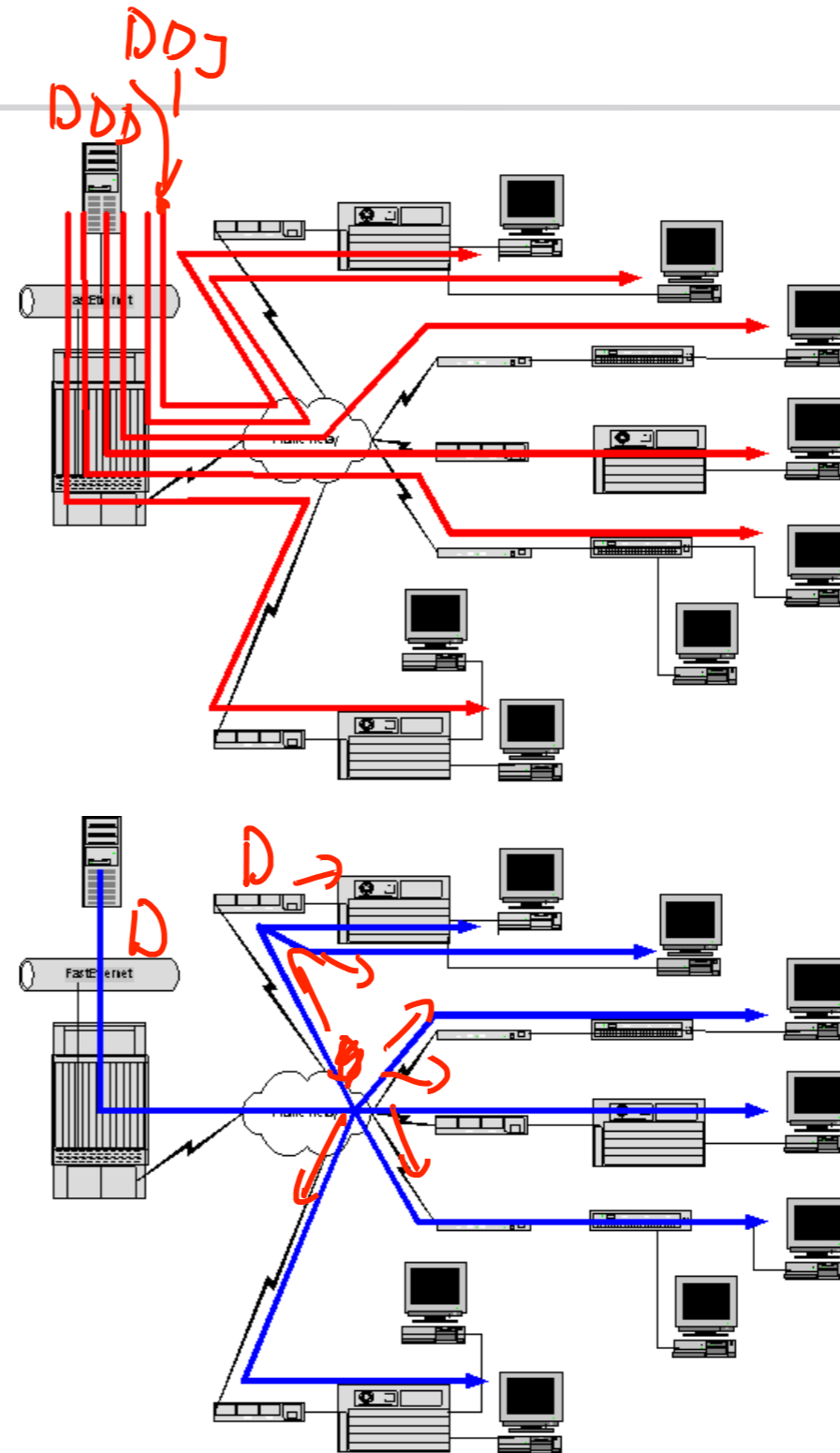
- Übertragung eines Stroms an viele Empfänger

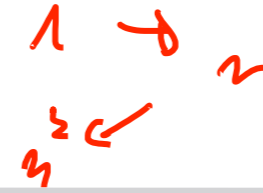
- Unicast

- Strom muss mehrfach einzeln übertragen werden
- Bottleneck am Sender

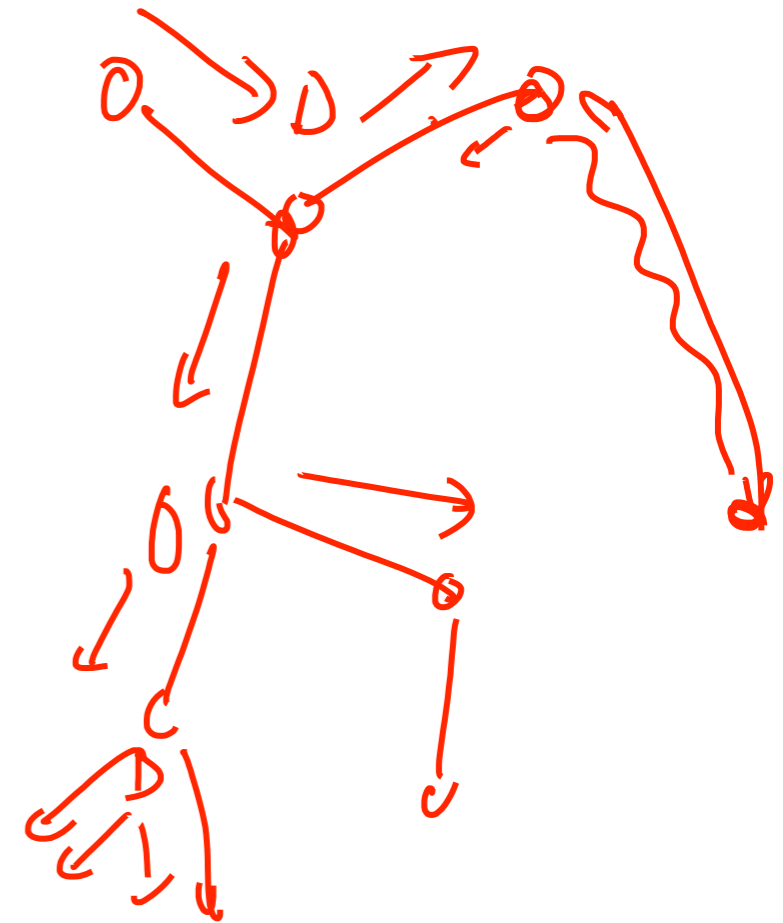
- Multicast

- Strom wird über die Router vervielfältigt
- Kein Bottleneck mehr

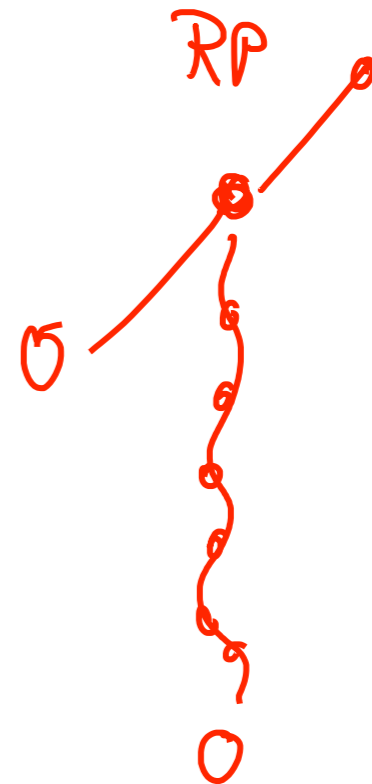




- IPv4 Multicast-Adressen
 - in der Klasse D (außerhalb des CIDR - Classless Interdomain Routings)
 - 224.0.0.0 - 239.255.255.255
 - in IPv6 mit Präfix FF
- Hosts melden sich per IGMP bei der Adresse an
 - IGMP = Internet Group Management Protocol
 - Nach der Anmeldung wird der Multicast-Tree aktualisiert
- Source sendet an die Multicast-Adresse
 - Router duplizieren die Nachrichten an den Routern
 - und verteilen sie in die Bäume
- Angemeldete Hosts erhalten diese Nachrichten
 - bis zu einem Time-Out
 - oder bis sie sich abmelden
- Achtung:
 - Kein TCP, nur UDP
 - Viele Router lehnen die Beförderung von Multicast-Nachrichten ab
 - Lösung: Tunneln



- Distance Vector Multicast Routing Protocol (DVMRP)
 - jahrelang eingesetzt in MBONE (insbesondere in Freiburg)
 - Eigene Routing-Tabelle für Multicast
- Protocol Independent Multicast (PIM)
 - im Sparse Mode (PIM-SM)
 - aktueller Standard
 - beschneidet den Multicast Baum
 - benutzt Unicast-Routing-Tabellen
 - ist damit weitestgehend protokollunabhängig
- Voraussetzung PIM-SM:
 - benötigt Rendezvous-Point (RP) in ein-Hop-Entfernung
 - RP muss PIM-SM unterstützen
 - oder Tunneling zu einem Proxy in der Nähe eines RP



Warum so wenig IP Multicast?

- Trotz erfolgreichen Einsatz
 - in Video-Übertragung von IETF-Meetings
 - MBONE (Multicast Backbone)
- gibt es wenig ISP welche IP Multicast in den Routern unterstützen
- ❖ Zusätzlicher Wartungsaufwand
 - Schwierig zu konfigurieren
 - Verschiedene Protokolle
- ❖ Gefahr von Denial-of-Service-Attacken
 - Implikationen größer als bei Unicast
- ❖ Transport-Protokoll
 - Nur UDP einsetzbar
 - Zuverlässige Protokolle
 - Vorwärtsfehlerkorrektur ✓
 - Oder proprietäre Protokolle in den Routern (z.B. CISCO)
- ❖ Marktsituation
 - Endkunden fragen kaum Multicast nach (benutzen lieber P2P-Netzwerke)
 - Wegen einzelner Dateien und weniger Abnehmer erscheint ein Multicast wenig erstrebenswert (Adressenknappheit!)

