

Vorlesung Kommunikationssysteme
Wintersemester 2024/25

TCP Überlastkontrolle,
Internet Routing und
Routingprotokolle

Christoph Lindemann

Comer Buch, Kapitel 25, 26

Zeitplan

Nr.	Datum	Thema
01	18.10.24	Organisation und Internet Trends
02	25.10.24	Programmierung mobiler Anwendungen mit Android
	01.11.24	Keine Vorlesung
03	08.11.24	Protokolldesign und das Internet
04	15.11.24	Anwendungen und Netzwerkprogrammierung
05	22.11.24	LAN und Medienzugriff
06	29.11.24	Ethernet und drahtlose Netze
07	06.12.24	LAN Komponenten und WAN Technologien
08	13.12.24	Internetworking und Adressierung mit IP
09	20.12.24	IP Datagramme
10	10.01.25	Zusätzliche Protokolle und Technologien
11	17.01.25	User Datagram Protocol und Transmission Control Protocol
12	24.01.25	TCP Überlastkontrolle / Internet Routing und Routingprotokolle
13	31.01.25	Ausblick: TCP für Hochgeschwindigkeitsnetze
14	07.02.25	Review der Vorlesung

Überblick

Ziele:

- ❑ Verständnis von Überlastkontrolle für den zuverlässigen Datentransport
- ❑ Verständnis für die Berechnung der Routen von IP Paketen und somit der Architektur des Internets

Themen:

- ❑ TCP
 - Überlastkontrolle
 - Segmentierung
- ❑ IP Routing
 - Routentabellen
 - Autonome Systeme
 - BGP
 - RIP
 - OSPF
 - Multicast Routing

TCP Überlastkontrolle

TCP Überlastkontrolle (1)

- ❑ Überlast ist Hauptgrund für Verspätung oder Verlust im Internet
- ❑ Retransmissions verschlimmern Überlast
- ❑ TCP misst Überlast anhand Änderungen des Delay
- ❑ Reduziert temporär Window Größe und damit Rate

TCP Überlastkontrolle (2)

□ TCP besitzt **Slow Start**

- Spezielle Überlastkontrolle für neue Verbindung oder bei verlorener Nachricht
- Am Anfang einzelne Nachricht mit Daten (Receiver Window wird nicht ganz gefüllt)
- Falls ACK ohne Verlust → doppelte Menge an Daten
- Exponentielle Steigerung bis zu Slow Start Threshold
- Danach lineare Steigerung der Datenmenge solange keine Überlast

TCP Überlastkontrolle (3)

- ❑ Nicht wirklich „slow“
 - RTT im Internet oft weniger als 100 ms → maximaler Durchsatz nach einer Sekunde
- ❑ Reagiert gut falls Netzwerk überlastet (vermeidet stärkere Überlastung)
- ❑ Schnelle Vermeidung von Congestion Collapse falls alle TCP Implementierungen den Regeln folgen

TCP Überlastkontrolle (4)

- ❑ Immer wieder kleinere Änderungen des Algorithmus
- ❑ Vorherige Erklärung entspricht *TAHOE*
- ❑ *TCP RENO* mit Fast Recovery (auch Fast Retransmit)
 - Höherer Durchsatz falls Verlust nur gelegentlich
- ❑ Zusätzlich: *TCP Vegas* und *NewReno* (meistgenutzt in 2000)
- ❑ Heute: *TCP CUBIC* und *BBR* (siehe auch VL 13)

SACK / ECN

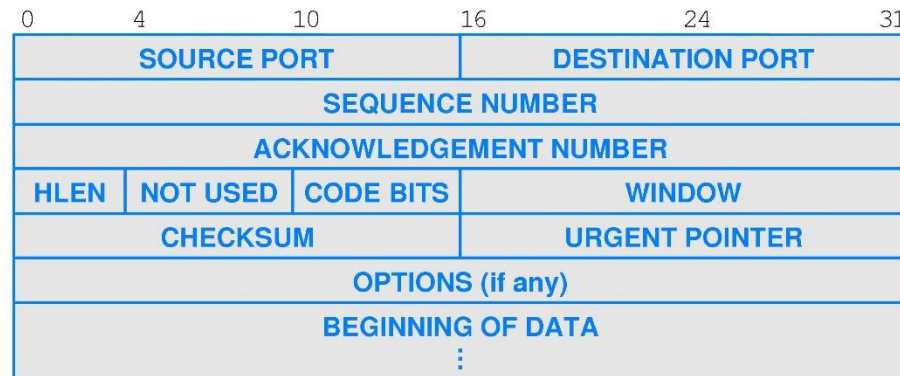
- ❑ TCP misst Überlast anhand Varianz der RTT
 - Betrachtet Netzwerk als Black Box und nutzt externe Maße
- ❑ **Selective Acknowledgement (SACK)**
 - Empfänger gibt an, welche Daten fehlen
 - Nur benötigte Daten werden neu versandt
 - Oft nicht nur zufällige Pakete verloren → Funktioniert nicht sehr gut
- ❑ **Explicit Congestion Notification (ECN)**
 - Router auf Pfad messen Überlast und markieren TCP Segment
 - Empfänger weiß von Überlast und teilt es in ACK mit
 - Dadurch aber Verzögerung: Sender muss auf ACK warten
 - Nur selten im Internet verwendet

TCP Segment Format (1)

- ❑ **TCP Segment:** Nachricht von TCP
 - Format einheitlich für Nachrichten mit Daten, ACK, 3-Way Handshake
- ❑ Felder einer Nachricht können sich auf beide Datenströme beziehen
 - Gleichzeitig Daten zu Empfänger, ACK für empfangene Pakete, Window Advertisement

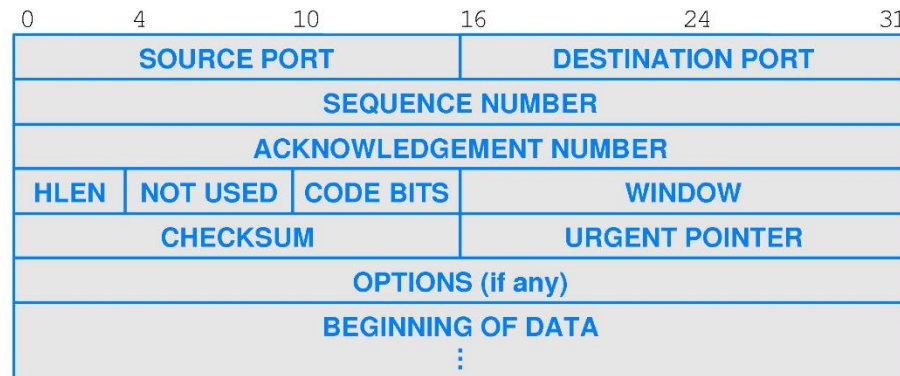
TCP Segment Format (2)

- ❑ ACKNOWLEDGEMENT NUMBER: Sequenznummer der erwarteten nächsten Daten
 - Gleiches ACK, falls ein Segment fehlt und weitere Segmente dahinter empfangen
- ❑ WINDOW: Freier Puffer



TCP Segment Format (2)

- ❑ SEQUENCE NUMBER: Sequenznummer des ersten Byte der Daten, Empfänger ordnet Segmente und berechnet ACKNOWLEDGEMENT NUMBER
- ❑ DESTINATION PORT: Programm auf Ziel, SOURCE PORT: Programm von Sender
- ❑ CHECKSUM: Berechnet für Header und Daten



Zusammenfassung

- ❑ TCP ist wichtigstes Transportprotokoll von TCP/IP
- ❑ Bietet Zuverlässigkeit, Flusskontrolle, Full Duplex Verbindungen, Stream Interface, Überlastkontrolle
- ❑ Garantiert Daten in richtiger Reihenfolge ohne Duplikate zu übertragen
- ❑ Alle Segmente benutzen gleiches Format (Kontroll- und Datenpakete)
- ❑ Sender überträgt verlorene Pakete anhand adaptiven Verfahren erneut

Internet Routing und Routingprotokolle

Einführung

- ❑ Weiterleitungstabellen in Routern müssen erstellt / aktualisiert werden
- ❑ Routing Informationen müssen an andere Systeme weitergeleitet werden
- ❑ Verschiedene Routingprotokolle existieren

Statisches / Dynamisches Routing

❑ Statisches Routing

- Weiterleitungstabelle bei Systemstart erstellt
- Kann nur manuell von Administrator geändert werden

❑ Dynamisches Routing

- **Route Propagation Software** aktualisiert Weiterleitungstabelle kontinuierlich
- Stellt sicher, dass Datagramme optimale Route nutzen
- Kommuniziert mit anderen Systemen, erkennt Netzwerkfehler
- Initialisierung mit Routen wie statisches Routing

Statisches Routing (1)

□ Vorteile

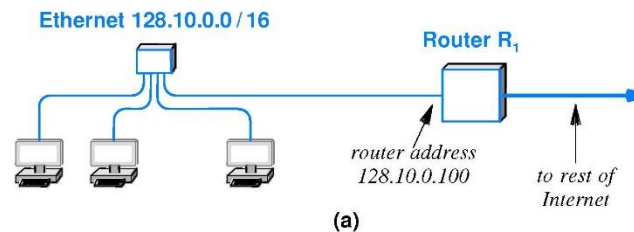
- Leicht zu definieren
- Benötigt keine weitere Software
- Kein zusätzlicher Verkehr im Internet
- Keine Rechenleistung für Weiterleitung der Informationen

□ Nachteile

- Unflexibel
- Netzwerkfehler oder Topologieänderungen nicht berücksichtigt

Statisches Routing (2)

- ❑ Großteil der Hosts nutzt statisches Routing
 - Oft eine Netzwerkverbindung und Router verbindet zu Internet
- ❑ Im Beispiel reichen schon zwei Einträge
- ❑ Nutzt Default Route für alle Hosts außerhalb des lokalen Netz



(b)

Net	Mask	Next hop
128.10.0.0	255.255.0.0	direct
default	0.0.0.0	128.10.0.100

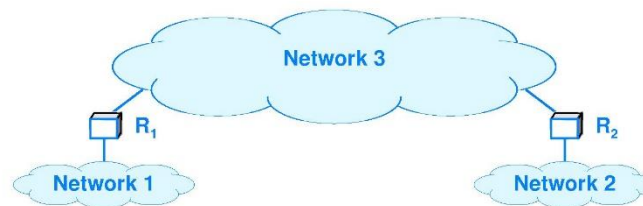
- (a) Typische Verbindung zu Internet
(b) Statische Weiterleitungstabelle eines Host

Dynamische Routen (1)

- ❑ Großteil der Router nutzt Dynamisches Routing
- ❑ Statisches Routing für Router R_1 in letzten Beispiel denkbar
 - Router hat Anbindung an ISP
 - Leitet alle Daten nur an diesen weiter
- ❑ Bei Verbindung von zwei ISPs müssen dynamische Routing Informationen ausgetauscht werden

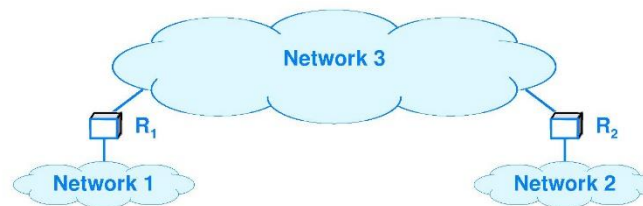
Dynamische Routen (2)

- ❑ Jeder Router kennt direkt verbundene Netze
- ❑ R_1 kennt nicht Netzwerk 2, R_2 kennt nicht Netzwerk 1
- ❑ Fügt ISP neues Netzwerk hinzu, muss Information durch Internet gereicht werden
- ❑ Netzwerkfehler und Überlastung mit manuellen Routen nicht sinnvoll zu umgehen



Dynamische Routen (3)

- Bei Verwendung von Route Propagation Software tauschen R_1 und R_2 Informationen über Netzwerke aus → Können Weiterleitungstabellen aktualisieren
- Falls Router R_2 abstürzt, erkennt dies Router R_1 und entfernt Routen zu Netzwerk 2
- Falls R_2 wieder erreichbar, wird Route wieder hinzugefügt



Routing im globalen Internet

- ❑ Zuviel Traffic falls Router im Internet allen anderen Routern Informationen mitteilt
- ❑ Internet nutzt deswegen hierarchisches Routing
- ❑ Router und Netzwerke sind in Gruppen organisiert
 - Router innerhalb einer Gruppe tauschen Informationen aus
 - Mindestens ein Router der Gruppe reicht gesammelte Informationen an andere Gruppen weiter
- ❑ Gruppen nicht in Größe beschränkt, Routingprotokoll in Gruppe nicht vorgeschrieben

Autonome Systeme

- ❑ **Autonomes System (AS):** Netzwerke und Router unter Kontrolle einer Organisation (Administrative Authority)
- ❑ Organisation kann z.B. ISP, Firma oder große Universität sein
- ❑ Große Organisation mit mehreren Standorten kann mehrere AS haben
- ❑ ISP kann sich in mehrere AS teilen
- ❑ Größe ökonomisch, technisch, administrativ bedingt
 - Firma mit mehreren AS verbunden zu regionalem ISP in jedem Land statt ein großes AS mit einer Verbindung zu Internet
 - Routingprotokolle erzeugen mehr Verkehr bei vielen Routern

Typen der Routingprotokolle

- Alle Routingprotokolle fallen in eine der Kategorien
 - Interior Gateway Protocols (IGPs)
 - Exterior Gateway Protocols (EGPs)

Interior Gateway Protocols (IGP)

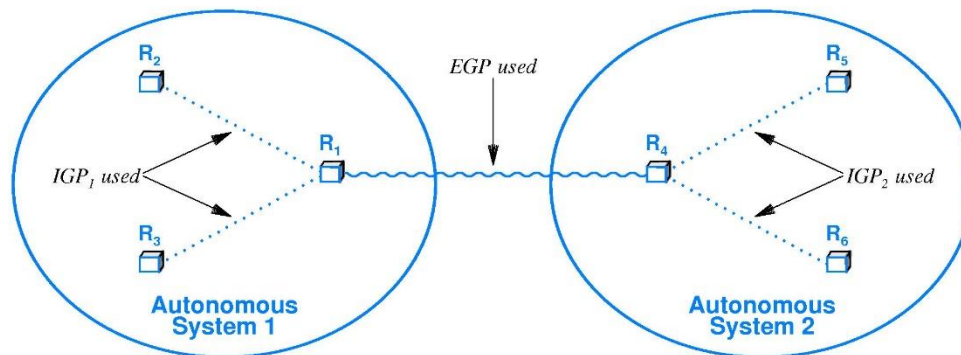
- ❑ Von Router innerhalb Autonomem System genutzt
- ❑ Jedes Autonome System kann ein eigenes IGP wählen
- ❑ Typischerweise einfach zu installieren und zu bedienen
- ❑ Kann Komplexität des Routing einschränken

Exterior Gateway Protocol (EGP)

- ❑ Router in AS nutzt EGP um mit Router in anderen AS zu kommunizieren
- ❑ EGP typischerweise komplexer zu installieren und zu warten
- ❑ Größere Flexibilität und weniger Overhead (Traffic)
- ❑ Routing Informationen von AS werden vor Übertragung gesammelt
- ❑ Information kann durch Richtlinien beschränkt werden (Policy Constraints)

IGP und EGP

- ❑ AS 1 nutzt IGP_1 : Alle Router in AS 1 kommunizieren damit
- ❑ AS 2 nutzt IGP_2 : Alle Router in AS 2 kommunizieren damit
- ❑ R_1 und R_4 kommunizieren zwischen AS mit EGP
- ❑ Sammeln Informationen vor Versand und propagieren an Router in eigenem AS



Optimale Routen, Metriken (1)

- ❑ Keine Übereinkunft was eine optimale Route ist
 - Remote Desktop: Geringer Delay
 - Download in Browser: Hoher Durchsatz
 - Audio Webcast: Wenig Jitter
- ❑ **Routing Metric:** Maß anhand dessen Routing Software Pfad wählt

Optimale Routen, Metriken (2)

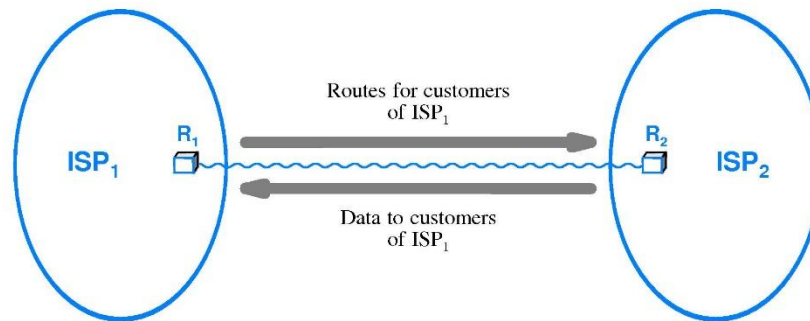
- ❑ Typischerweise Kombination von administrativen Kosten und Hop Count genutzt
- ❑ Hop gibt Anzahl Router bzw. Netzwerke auf Weg zu Ziel an
- ❑ Administrative Kosten manuell zugeteilt
 - Kommunikation von Accounting und Gehaltsstelle in Firma nicht durch Kundennetz leiten → Manuell höhere Kosten für den Pfad

Optimale Routen, Metriken (3)

- ❑ Nur IGPs nutzen Routingmetriken
- ❑ Jedes AS wählt eigene Routingmetrik und bestimmt danach Pfade
- ❑ EGP versucht einen Pfad zu finden, aber nicht optimalen Pfad
 - Vergleich wäre nicht möglich, da jedes AS mit eigener Metrik
 - Beispiel: Ein AS misst Hops, das andere aber Durchsatz

Routen und Datenverkehr

- Datenverkehr für ein Ziel erfolgt in entgegengesetzter Richtung des Routenverkehr
- Bevor Netzwerk in ISP_1 Daten bekommt, muss ISP_1 Routen nach außen mitteilen



BGP (1)

- ❑ **Border Gateway Protocol (BGP):** DAS Exterior Gateway Protocol des Internet
- ❑ Seit langem vierte Version BGP-4 genutzt
- ❑ Routing unter Autonomen Systemen:
 - Routen als Pfade von AS gegeben, keine Details über Router innerhalb AS
- ❑ Richtlinien:
 - Sender und Empfänger können Richtlinien haben (z.B. Einschränkung der propagierten Routen)

BGP (2)

❑ Transit Routing

- AS ist Transit System, falls es Daten zu anderen AS durchleitet, sonst Stub System
- AS kann Transit Verkehr verweigern

❑ Zuverlässiger Transport:

- Nutzt TCP für Kommunikation

❑ BGP ist das EGP, welches von Tier-1 ISPs genutzt wird

RIP (1)

- ❑ **Routing Information Protocol (RIP):** Eins der ersten IGP im Internet, erlaubt Routing innerhalb AS
- ❑ **Hop Count Metrik:** Distanz in Netzwerk Hops gemessen, direkt verbundenes Netzwerk zählt als 1 Hop
- ❑ **Unzuverlässiger Transport:** Nutzt UDP für Transfer
- ❑ **Broadcast oder Multicast:** Für LAN Technologien mit Broadcast oder Multicast gedacht, Multicast ab Version 2

RIP (2)

- ❑ IPv4 CIDR und Subnetting: Adressmaske zu Zieladresse ab Version 2
- ❑ Unterstützung von Default Routen
- ❑ Nutzt Distance Vector Algorithmus
- ❑ Passive Version für Hosts: Host kann Informationen abhören und eigene Weiterleitungstabelle aktualisieren, Nützlich falls mehrere Router in Netzwerk
- ❑ Erweiterung für IPv6: RIP next generation (RIPng)

RIP (3)

- ❑ Ausgehende Nachricht enthält erreichbare Netzwerke mit Distanz
- ❑ Jeder Eintrag als (Zielnetzwerk, Distanz) gegeben mit Distanz in Hops
- ❑ Bei Empfang von Nachricht wird Weiterleitungstabelle aktualisiert
 - Falls Netzwerk bisher nicht erreichbar oder längere Distanz, wird Route durch Route zu Sender ersetzt
- ❑ Hauptvorteil von RIP ist Einfachheit
- ❑ Default Route wird propagiert (z.B. Route zu ISP)

RIP (4)

- Jeder Eintrag enthält IP-Adresse mit Adressmaske, Nächsten Hop, Distanz

0	8	16	24	31
COMMAND (1-5)		VERSION (2)		MUST BE ZERO
FAMILY OF NET 1		ROUTE TAG FOR NET 1		
IP ADDRESS OF NET 1				
ADDRESS MASK FOR NET 1				
NEXT HOP FOR NET 1				
DISTANCE TO NET 1				
FAMILY OF NET 2		ROUTE TAG FOR NET 2		
IP ADDRESS OF NET 2				
ADDRESS MASK FOR NET 2				
NEXT HOP FOR NET 2				
DISTANCE TO NET 2				
...				

Format von RIP Version 2 Update Nachricht

OSPF (1)

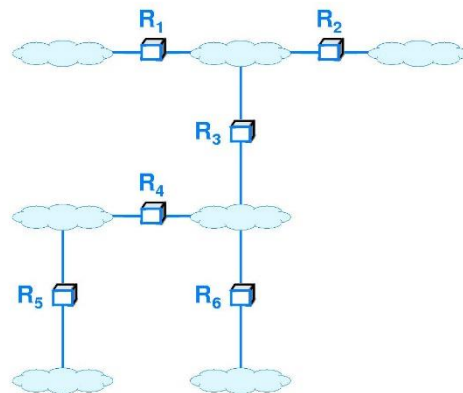
- ❑ Anzahl Nachrichten von RIP proportional zu Anzahl erreichbarer Netzwerke → Verzögerung, viele Berechnungen
- ❑ Funktioniert nur für wenige Router
- ❑ **Open Shortest Path First Protocol (OSPF):** IGP, welches Dijkstra Algorithmus nutzt
- ❑ CIDR Unterstützung: Adressmaske zu jeder IP Adresse
- ❑ Authentifizierter Nachrichtenaustausch
- ❑ Andere Routen können importiert werden (z.B. von BGP)

OSPF (2)

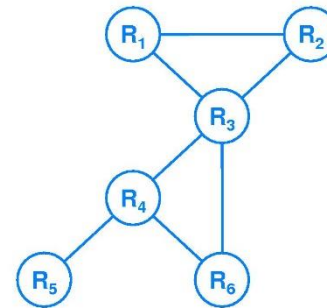
- ❑ Link-State Algorithmus wird verwendet
- ❑ Metriken werden unterstützt (Administrator kann Kosten zu Route zuweisen)
- ❑ OSPFv3 hat Unterstützung für IPv6
- ❑ Unterstützung von Multi-Access Netzwerken:
 - Link-State Routing z.B. in Ethernet ineffizient, da alle Router Link Status broadcasten
 - OSPF bestimmt speziellen Router dafür

OSPF Graph

- ❑ Link-State Routing stellt Netzwerk als Graph dar (Knoten ist Router, Kante ist Verbindung)
- ❑ Jedes Paar von Routern testet Verbindung und broadcastet Link-State Nachricht an alle Router
- ❑ Jeder Router aktualisiert lokale Kopie des Graphen



(a)



(b)

(a) Beispiel Topologie

(b) Zugehöriger OSPF Graph

OSPF Areas

- ❑ OSPF unterstützt hierarchisches Routing
- ❑ AS kann für Routing in Areas weiter unterteilt werden
- ❑ Router kennt Grenze der Area (alle anderen Router der Area)
- ❑ Router tauschen nur Link-State Nachrichten innerhalb Area aus
- ❑ Ein Router jeder Area kommuniziert mit Routern in anderen Areas
 - Sammeln dazu Routing Informationen ihrer Area
 - Skaliert zu viel größeren AS als andere Routingprotokolle

IS-IS (1)

- ❑ Intermediate System - Intermediate System (IS-IS)
- ❑ Weiteres IGP, nutzt auch Link-State Nachrichten und Dijkstra Algorithmus wie OSPF
- ❑ Unterschiede zu OSPF
 - Anfangs proprietär (von DEC), OSPF als offener Standard erstellt
 - OSPF für IP konstruiert, IS-IS für CLNS (Netzwerkprotokoll des OSI Protokollstack)

IS-IS (2)

- ❑ Anfangs OSPF viel populärer, wurde erweitert und dadurch immer komplexer
- ❑ IS-IS nach Auflösung von DEC frei verfügbar
- ❑ Neu entwickelte Version integriert IP und hat IPv6 Unterstützung
- ❑ Verwendung in großen ISPs aufgrund weniger Overhead als OSPF attraktiv

IP Multicast (1)

- ❑ Applikation kann jederzeit Gruppe beitreten → erhält alle Pakete der Gruppe
 - Host der Applikation informiert Router in der Nähe
- ❑ Nur ein Paket an Host, falls mehrere Applikationen auf Host in selber Gruppe

IP Multicast (2)

- ❑ Kann jederzeit Gruppe verlassen
 - Host sendet periodisch Group Membership Nachrichten zu Router
 - Informiert Router, wenn keine Applikation des Hosts mehr zu Gruppe gehört
- ❑ Sender / Empfänger kennen nicht Anzahl und Identität der Gruppenmitglieder
- ❑ Sender muss nicht Teil der Gruppe sein, beliebige Applikation kann an Gruppe senden

Multicast Protokolle

- ❑ Es existiert kein Multicast Routing für gesamtes Internet
- ❑ Mehrere Vorschläge existieren
 - Distance Vector Multicast Routing Protocol (DVMRP)
 - Core Based Trees (CBT)
 - Protocol Independent Multicast - Sparse Mode (PIM-SM)
 - Protocol Independent Multicast - Dense Mode (PIM-DM)
 - Multicast Extensions To The Open Shortest Path First Protocol (MOSPF)

Protocol	Type
DVMRP	Configuration-And-Tunneling
CBT	Core-Based-Discovery
PIM-SM	Core-Based-Discovery
PIM-DM	Flood-And-Prune
MOSPF	Link-State (within an organization)

Zusammenfassung

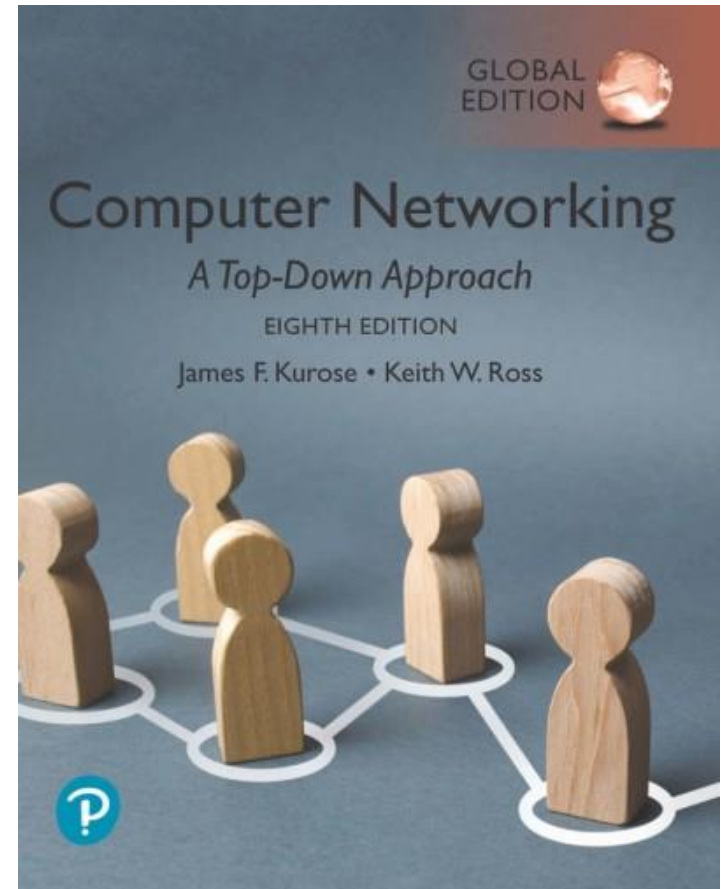
- ❑ Großteil der Hosts nutzt statisches Routing
- ❑ Großteil der Router nutzt dynamisches Routing → Aktualisieren Weiterleitungstabelle kontinuierlich
- ❑ Internet ist in Autonome Systeme geteilt
- ❑ Routing Protokolle werden in EGP und IGP unterschieden
 - BGP ist das primäre EGP des Internet
 - Mögliche IGP: RIP, OSPF, IS-IS
- ❑ Kein Multicast Routing existiert für gesamtes Internet

Weiterführendes Lehrbuch zur Vorlesung

James Kurose, Keith Ross,
Computer Networking: A Top-
Down Approach, Global
Edition, 8. Auflage, Pearson,
2021.

An Uni als E-Book

<https://katalog.ub.uni-leipzig.de/Record/O-1771738375>



Selbststudium

Zum Vertiefen der Inhalte dieser Vorlesung

Leseaufgabe zum Selbststudium bis 7.2.2025:

James Kurose, Keith Ross, Computer Networking: A Top-Down Approach, Pearson, 2021. S. 211 - 331
Chapter 3: The Transport Layer.

Klausur

Termin:

27.02.2025, 10:30 Uhr bis 11:30 Uhr

Ort:

Audimax, Augusteum

Viel Erfolg!!