Netzwerktechnologien
3 VO

Dr. Ivan Gojmerac
ivan.gojmerac@univie.ac.at

7. Vorlesungseinheit, 08. Mai 2013

Bachelorstudium Medieninformatik
SS 2013

4.1.1 Verbindungsauf- und Verbindungsabbau

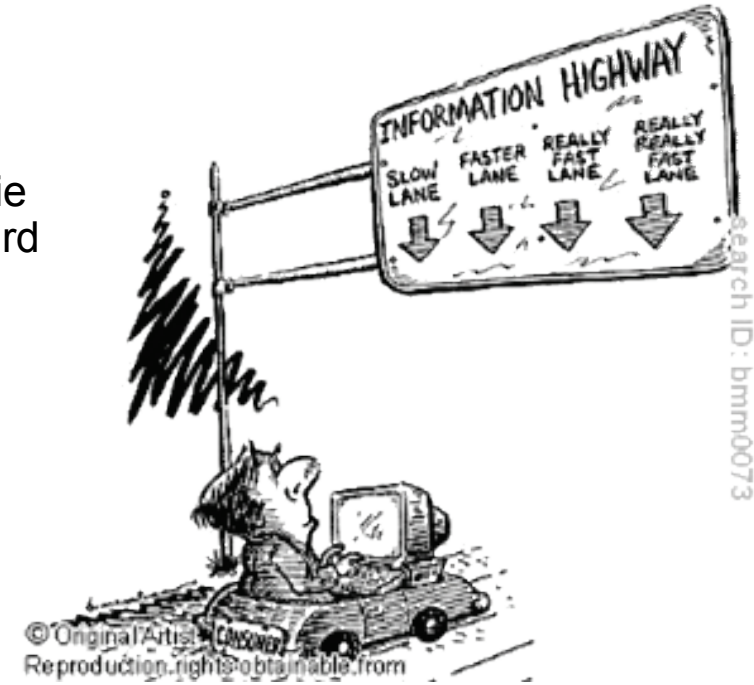
- Dritte wichtige Funktion in einigen Netzwerkarchitekturen:
  - ATM (Asynchronous Transfer Mode), Frame Relay, X.25
- Bevor Daten übertragen werden können, wird eine virtuelle Verbindung vom Sender zum Empfänger aufgebaut
  - Nach der Übertragung wird diese Verbindung wieder abgebaut
- Unterschied Verbindungen auf der Transportschicht und der Netzwerkschicht:
  - *Netzwerkschicht*: zwischen zwei Hosts (dazwischenliegende Router können involviert sein)
  - *Transportschicht*: zwischen zwei Prozessen

4.1.2 Dienstmodelle der Netzwerkschicht

Unterschiedliche Anforderungen verlangen unterschiedliche Dienstmodelle:

- Beispiele für einzelne Datagramme:
  - Garantierte Zustellung
  - Garantierte Zustellung in weniger als 40 ms

- Beispiele für einen Fluss von Datagrammen:
  - Reihenfolgeerhaltende Auslieferung
  - Garantierte minimale Datenrate
  - Beschränkung der Schwankungen in der Zeit, die für den Transport von Datagrammen benötigt wird



4.1.2 Dienstmodelle der Netzwerkschicht

Netzwerk-architektur	Dienst-modell	Band-breiten-garantien	Garantie der Ver-lustfreiheit	Reihen-folge	Zeit-garantien	Hinweis auf Überlast
Internet	Best Effort	keine	nein	beliebige möglich	nicht unterstützt	keiner
ATM	CBR	garantiert eine kon-stante Rate	ja	in korrekter Reihenfolge	unterstützt	Überlast tritt nicht auf
ATM	ABR	garantier-tes Mini-mum	nein	in korrekter Reihenfolge	nicht unterstützt	Überlasthin-weise werden verwendet

4.2 Virtuelle Leitungen und Datagrammnetzwerke

Verbindungsorientierter und verbindungsloser Netzwerkschichtdienst

- Ein Datagrammnetzwerk verwendet eine verbindungslose Netzwerkschicht
- Ein Netzwerk mit virtuellen Leitungen verwendet eine verbindungsorientierte Netzwerkschicht
- Analog zur Transportschicht. Aber:
  - ! Dienst: Host-zu-Host (nicht Prozess-zu-Prozess)
  - ! Keine Wahlmöglichkeit: Ein Netzwerk bietet das eine oder das andere an
  - ! Implementierung: im Inneren des Netzwerkes

→ Achtung! In höheren Schichten kann immer noch ein verbindungsorientierter Dienst (z.B. TCP) über eine verbindungslose Vermittlungsschicht (z.B. IP) realisiert werden!

4.2.1 Netzwerke mit virtuellen Leitungen

→ Der Pfad zwischen Sender und Empfänger verhält sich wie eine Telefonleitung

- Aufbau einer Verbindung, bevor Daten transportiert werden können
  Danach: Abbau der Verbindung
- Jedes Paket beinhaltet einen VC-Identifier (Virtual Channel = Virtuelle Leitung) und keine Zieladresse
- *Jeder Router* auf dem Pfad vom Sender zum Empfänger verwaltet einen Zustand für diese Verbindung
- Ressourcen von Links und des Routers können einer virtuellen Leitung zugeordnet sein (zugeordnete Ressourcen = vorhersagbare Dienstgüte)

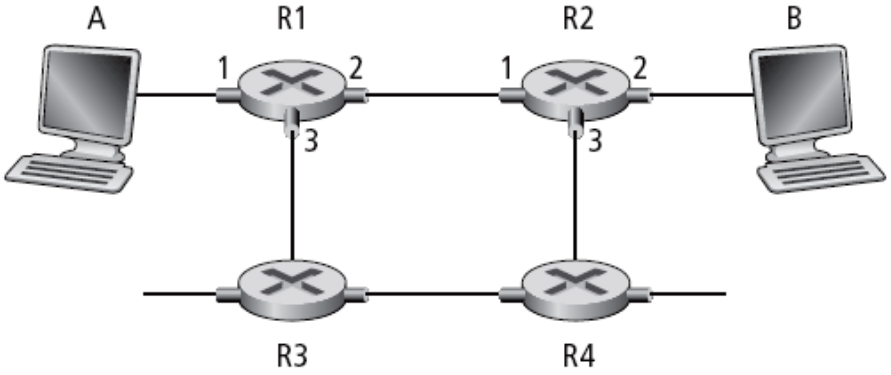
4.2.1 Implementierung virtueller Leitungen

Eine virtuelle Leitung besteht aus:

1. Einem Pfad vom Sender zum Empfänger
2. VC-Identifier, ein Identifier für jeden Link entlang des Pfades
3. Einträgen in den Weiterleitungstabellen der Router auf dem Pfad

- Pakete, die zu einem VC gehören, sind durch einen VC-Identifier (nicht durch die Zieladresse!) gekennzeichnet
- Der VC-Identifier desselben Paketes kann von Link zu Link unterschiedlich sein
  - Die neuen VC-Identifier stehen in der Weiterleitungstabelle der Router

4.2.1 Implementierung virtueller Leitungen

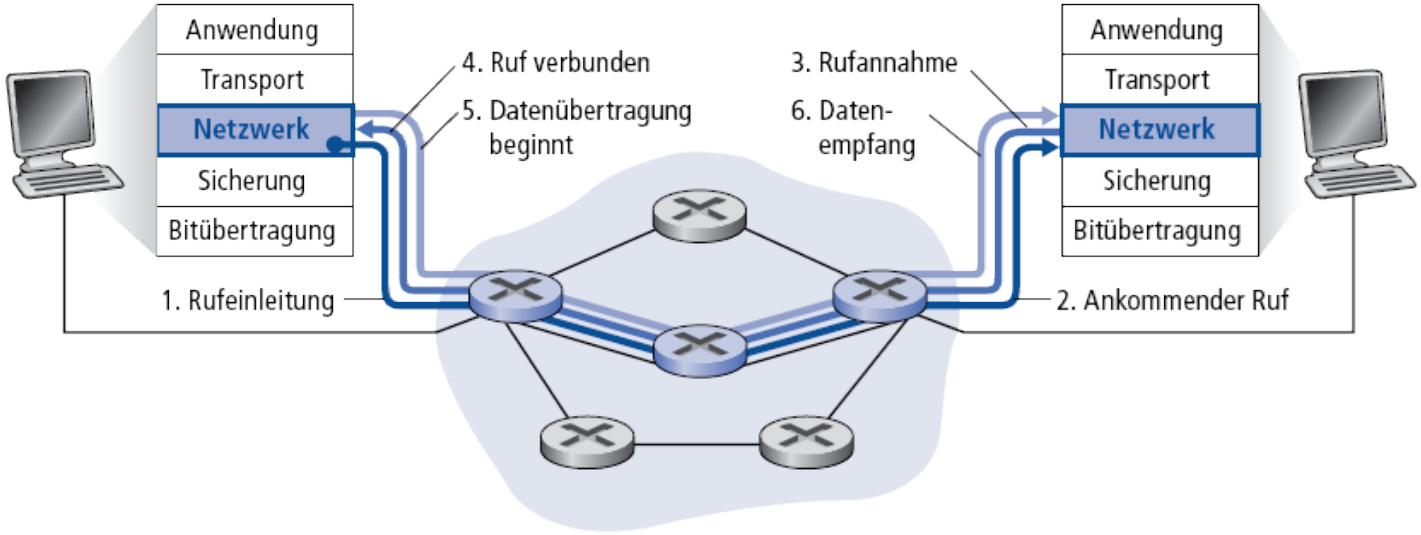


Weiterleitungstabelle in R1:

Eingehende Schnittstelle	Eingehende VC-Nummer	Ausgehende Schnittstelle	Ausgehende VC-Nummer
1	12	2	22
2	63	1	18
3	7	2	17
1	97	3	87
...	...	...	...

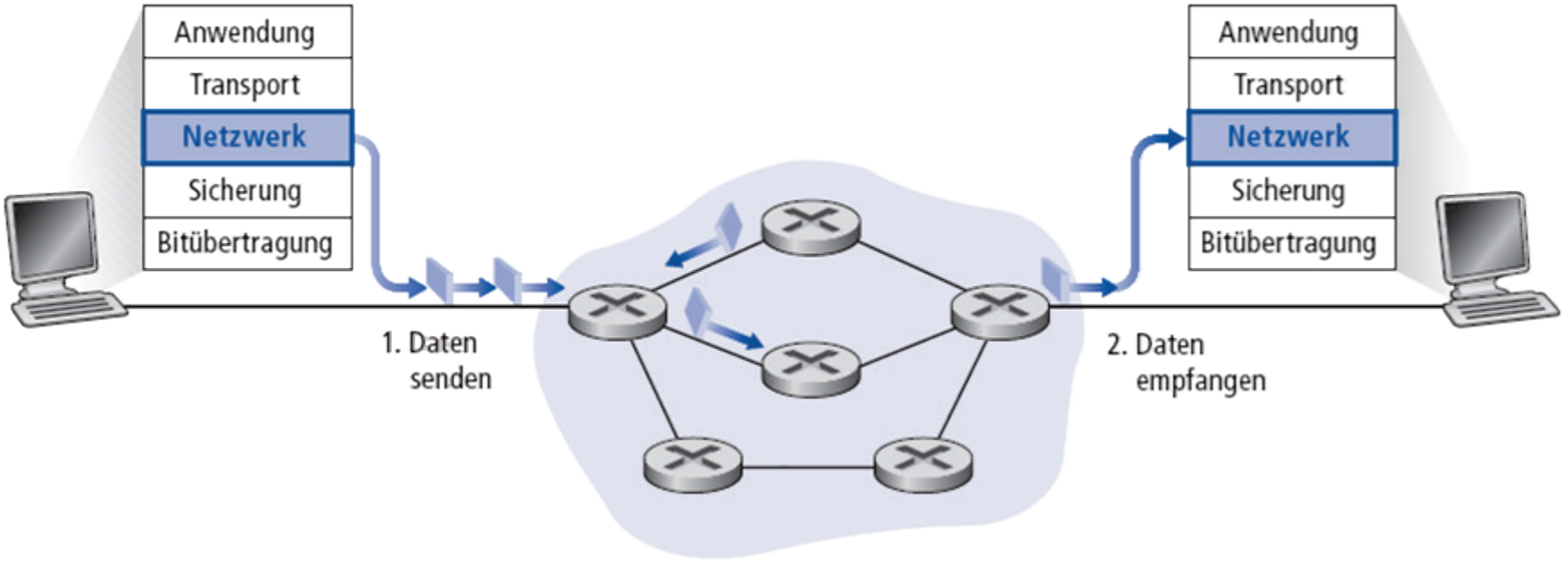
→ Router verwalten Zustand für jede Verbindung!

4.2.1 Virtuelle Leitungen: Signalisierungsprotokolle



- Verwendet zum Aufbau, Aufrechterhalten und Abbau von virtuellen Leitungen
- Verwendet in ATM, Frame Relay und X.25
- Nicht im Internet!

4.2.2 Datagrammnetzwerke



- Kein Verbindungsaufbau auf der Netzwerkschicht
- Router halten keinen Zustand für Ende-zu-Ende-Verbindungen
  - Auf Netzwerkebene gibt es das Konzept einer “Verbindung” nicht!
- Pakete werden unter Verwendung einer Zieladresse weitergeleitet
  - Pakete für dasselbe Sender-Empfänger-Paar können je nach Konfiguration in unterschiedlichen Richtungen einen unterschiedlichen Pfad nehmen

4.2.2 Weiterleitungstabelle eines Routers

Zieladressbereich	Schnittstelle
11001000 00010111 00010000 00000000	
bis	0
11001000 00010111 00010111 11111111	
11001000 00010111 00011000 00000000	
bis	1
11001000 00010111 00011000 11111111	
11001000 00010111 00011001 00000000	
bis	2
11001000 00010111 00011111 11111111	
sonst	3

4.2.2 Longest Prefix Matching

Passender Präfix	Schnittstelle
11001000 00010111 00010	0
11001000 0 0010111 00011000	1
11001000 00010111 00011	2
sonst	3

Beispiele

Adresse: 11001000 00010111 00010110 10100001 → **Schnittstelle 0**

Passt zum 3. Eintrag! → Schnittstelle 2

Adresse: 11001000 00010111 00011000 10101010

Passt zum 2. Eintrag! → **Schnittstelle 1**

Gibt es mehrere Treffer, verwendet der Router die **Longest-Prefix-Matching-Regel** (längstes übereinstimmendes Präfix).

4.2.2 Datagramme und virtuelle Verbindungen

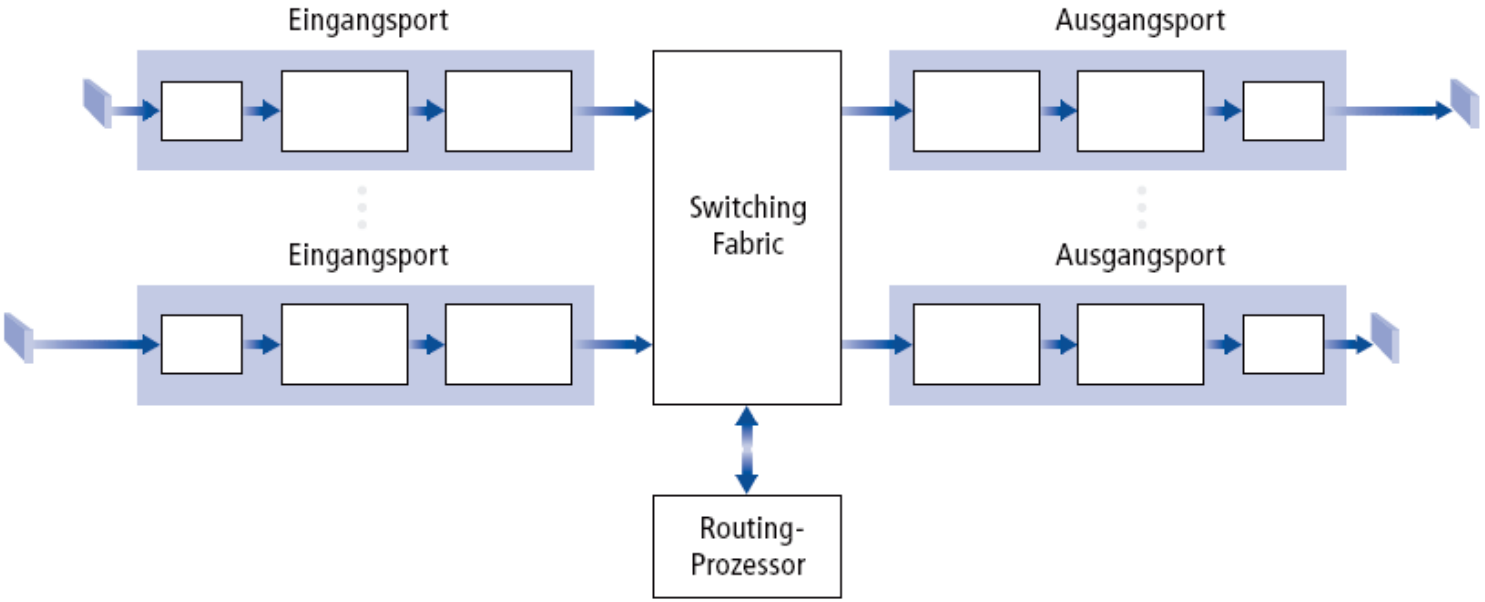
Internet

- Datenaustausch zwischen Computern
  - Keine Echtzeitanforderungen
- Mächtige Endsysteme
  - Können sich anpassen und Fehler beheben
  - Konsequenz: einfaches Netzwerk, Komplexität in den Endsystemen
- Vielzahl verschiedener Links
  - Unterschiedliche Charakteristika
  - Einheitlicher Dienst schwierig zu realisieren

ATM (Asynchronous Transfer Mode)

- Stammt von der klassischen Telefontechnologie ab
- Menschliche Kommunikation
  - Hohe Anforderungen an Echtzeit und Zuverlässigkeit
  - Dienstgarantien sind notwendig
- Einfache Endsysteme
  - Telefone
  - Konsequenz: einfache Endsysteme, Komplexität im Netzwerk

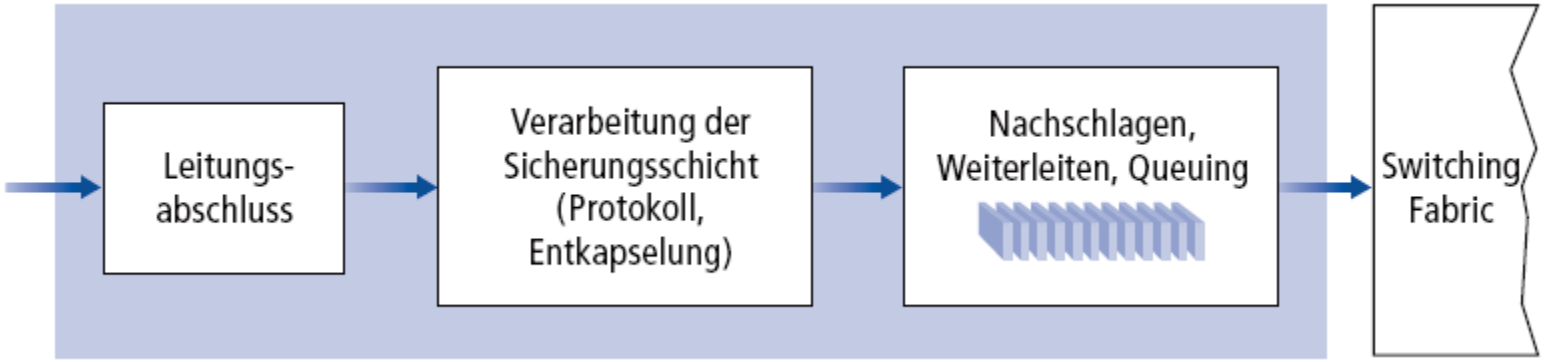
4.3 Router



Zwei wichtige Aufgaben eines Routers:

- Ausführen von Routing-Algorithmen und -Protokollen
  - RIP, OSPF, BGP
- Weiterleiten von Datagrammen von einem eingehenden zu einem ausgehenden Link

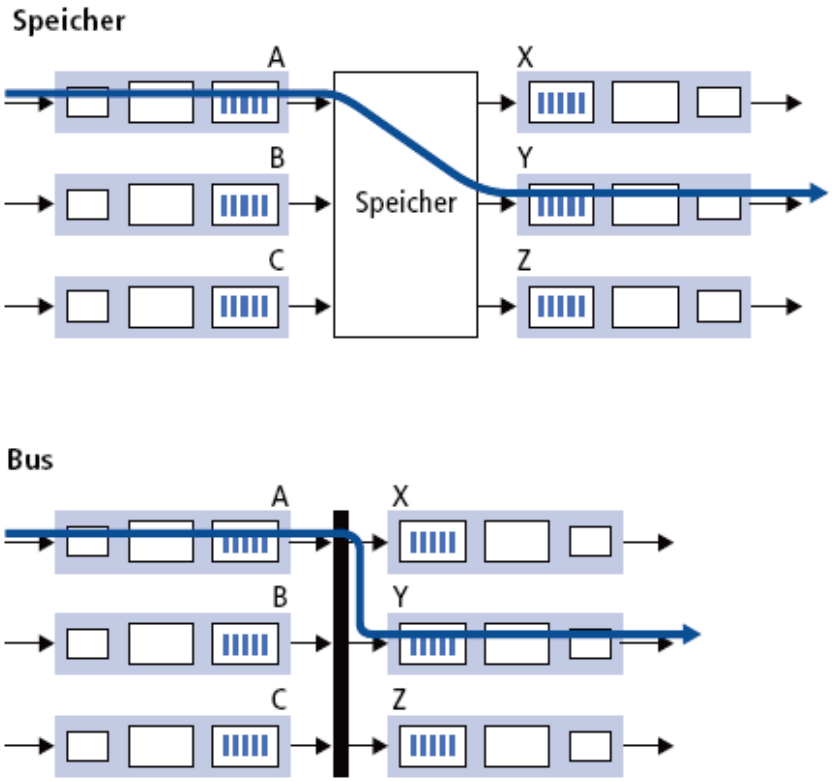
4.3.1 Eingangsports



- Leitungsabschluss: physikalische Schicht, Bits empfangen
- Sicherungsschicht: z.B. Ethernet (s. Kapitel 5)
- Nachschlagen, Weiterleiten, Queuing:
  - Suche nach einem geeigneten Ausgangsport
  - Dezentral, Kopie der Routing-Tabelle (oder Teile davon) notwendig
  - Ziel: Behandlung der Pakete mit „line speed“, also mit der Geschwindigkeit der Eingangsleitung des Ports
  - Puffern von Paketen, wenn die Switching Fabric belegt ist

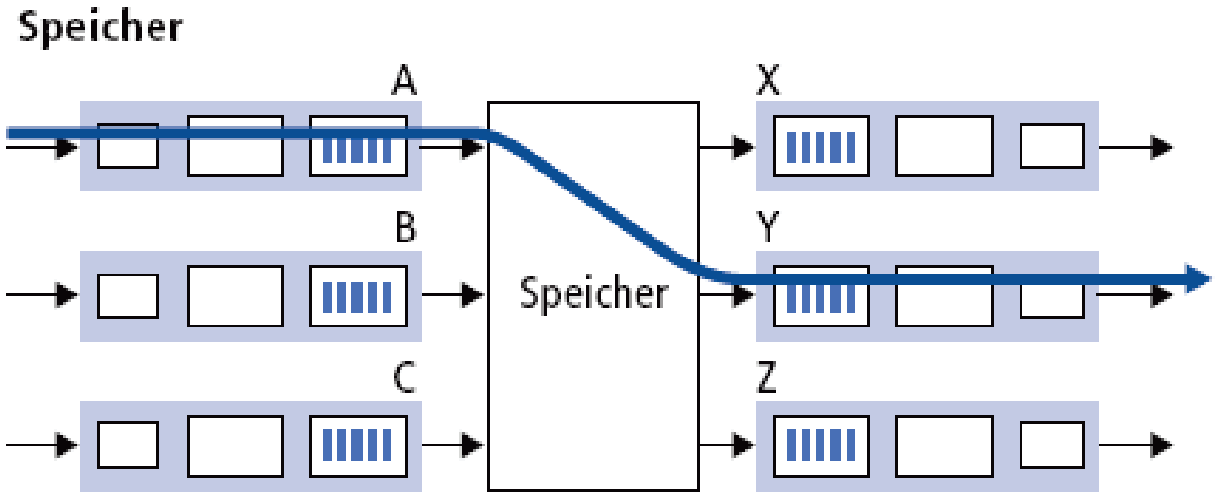
4.3.2 Das Switching Fabric

Drei Switching-Techniken:



4.3.2 Das Switching Fabric

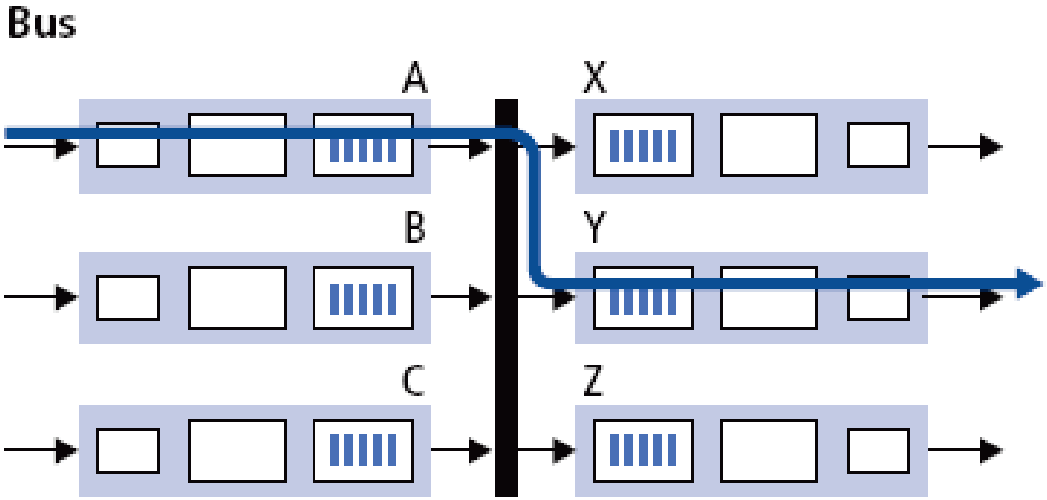
1. Switching über den Speicher:



- Erste Routergeneration:
  - „Normale“ Rechner, Switching wird über die CPU durchgeführt
  - Paket von Eingangsport in den Hauptspeicher kopieren
  - Paket vom Hauptspeicher in den Ausgangsport kopieren
  - Geschwindigkeit durch Speicherbus beschränkt!
    - Zwei Speicherzugriffe: einer zum Schreiben, einer zum Lesen

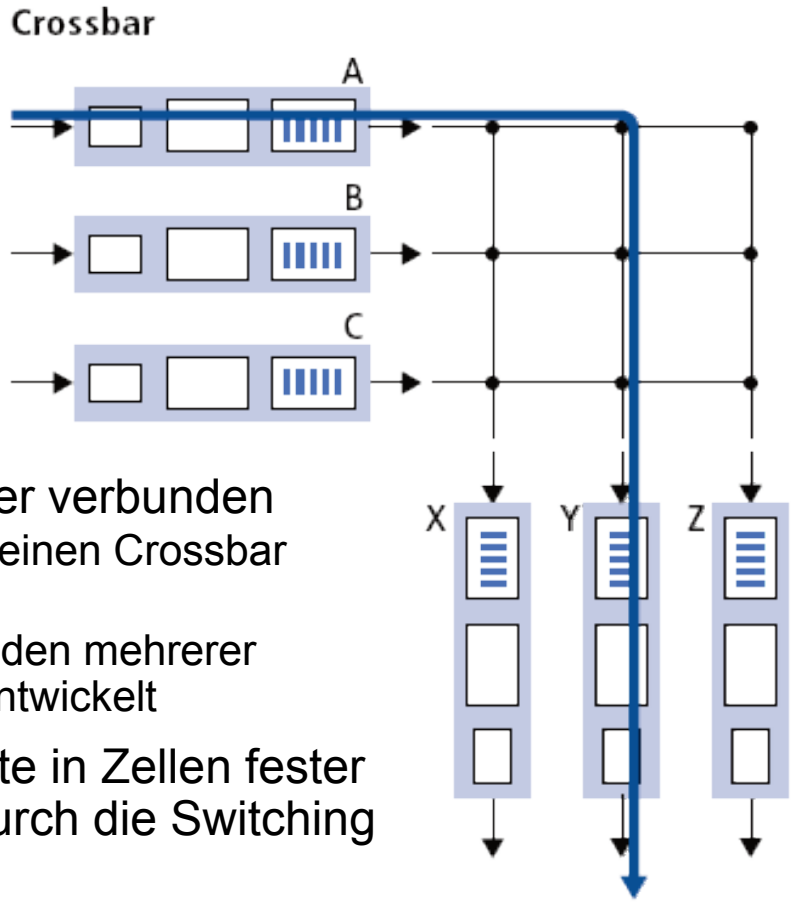
4.3.2 Das Switching Fabric

2. Switching über einen Bus:



- Alle Ports teilen sich einen gemeinsamen Bus
- **Bus Contention**: Die gesamte Kommunikation erfolgt über den Bus, dieser beschränkt die Bandbreite des Routers
  - Aber: nur eine Busoperation (nicht zwei!)
  - *Beispiel*: 32-Gbps-Bus, Cisco 5600, ausreichend für Zugangsrouter und Router für Firmennetze (nicht geeignet im Backbone!)

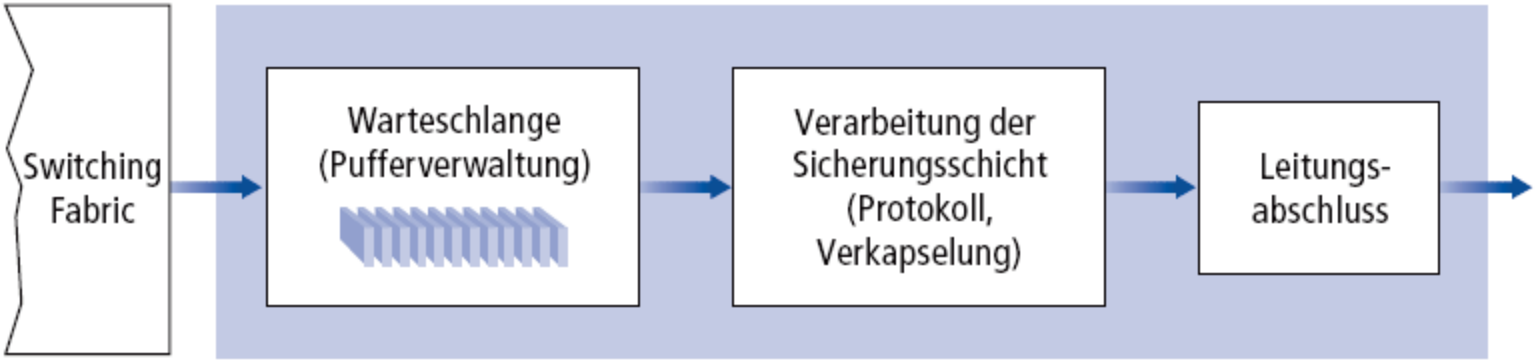
4.3.2 Das Switching Fabric



3. Switching über ein Spezialnetz:

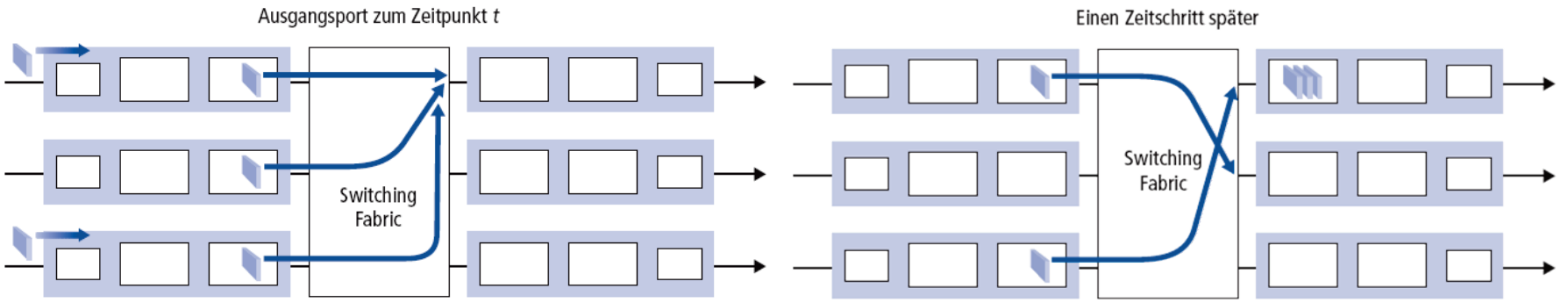
- Ports sind über ein Netzwerk miteinander verbunden
  - Beispielsweise alle Eingangsports über einen Crossbar mit allen Ausgangsports
  - Technologie ursprünglich für das Verbinden mehrerer Prozessoren in einem Parallelrechner entwickelt
- Weitere Fortschritte: Zerlegen der Pakete in Zellen fester Größe, Zellen können dann schneller durch die Switching Fabric geleitet werden
  - *Beispiel*: Cisco 12000, Switching von 60 Gbps durch das interne Netz

4.3.3 Ausgangsports



- Prinzipiell: analog zum Eingangsport!
- Einfacher, da die Entscheidung über die Weiterleitung schon getroffen ist

4.3.4 Puffern im Ausgangsport



- Puffern von Paketen, wenn sie schneller aus der Switching Fabric kommen, als sie auf die Leitung gelegt werden können
- Auswirkungen:
  - Gepufferte Pakete werden verzögert
  - Wenn der Puffer überläuft, müssen Pakete verworfen werden
- **„Scheduling Discipline“:** bestimmt die Reihenfolge, in der gepufferte Pakete auf die Leitung gelegt werden

4.3.4 Puffern im Ausgangsport

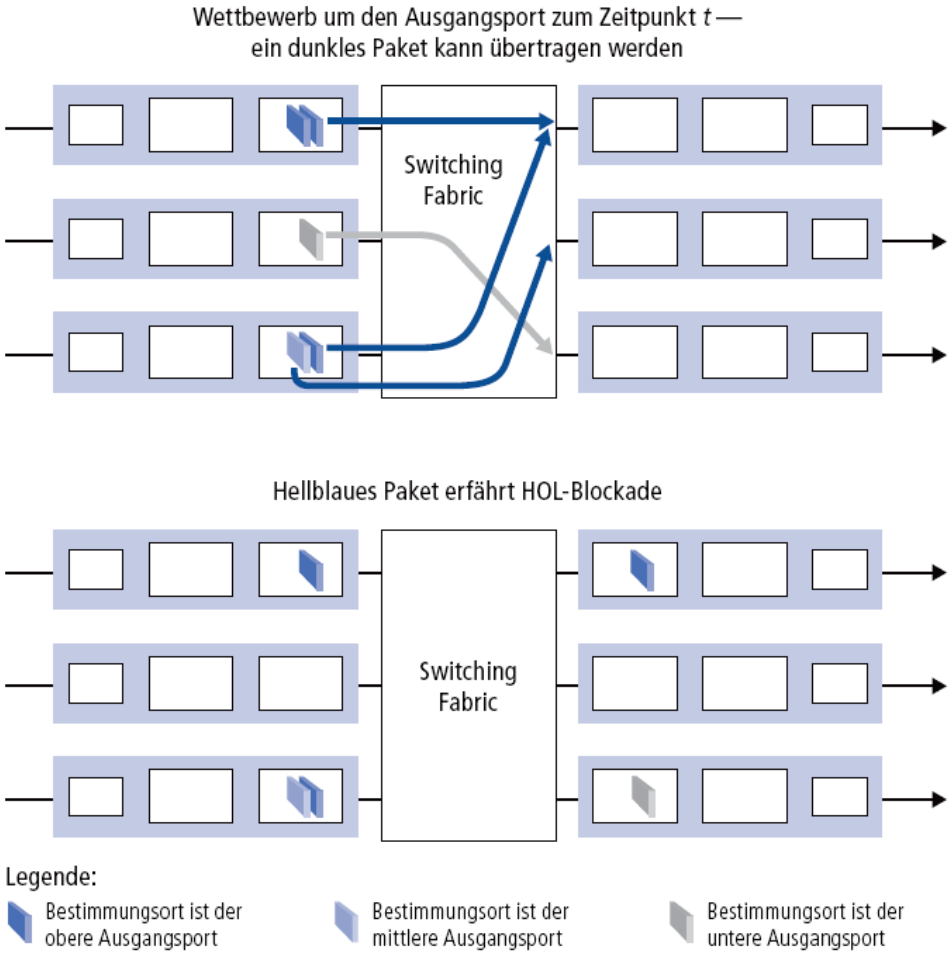
Wie groß sollten die Puffer sein?

- RFC 3439 beschreibt folgende **Faustregel**: Die Größe des Puffers sollte der Rundlaufzeit (RTT, z.B. 250 ms) multipliziert mit der Datenrate des Links entsprechen
  - Bei einem 10 Gps Link, 250 ms RTT, ergibt das 2,5 Gbit Puffer
- Neuere Empfehlungen: bei N Datenflüssen und Link-Datenrate C:

$$\frac{RTT \cdot C}{\sqrt{N}}$$

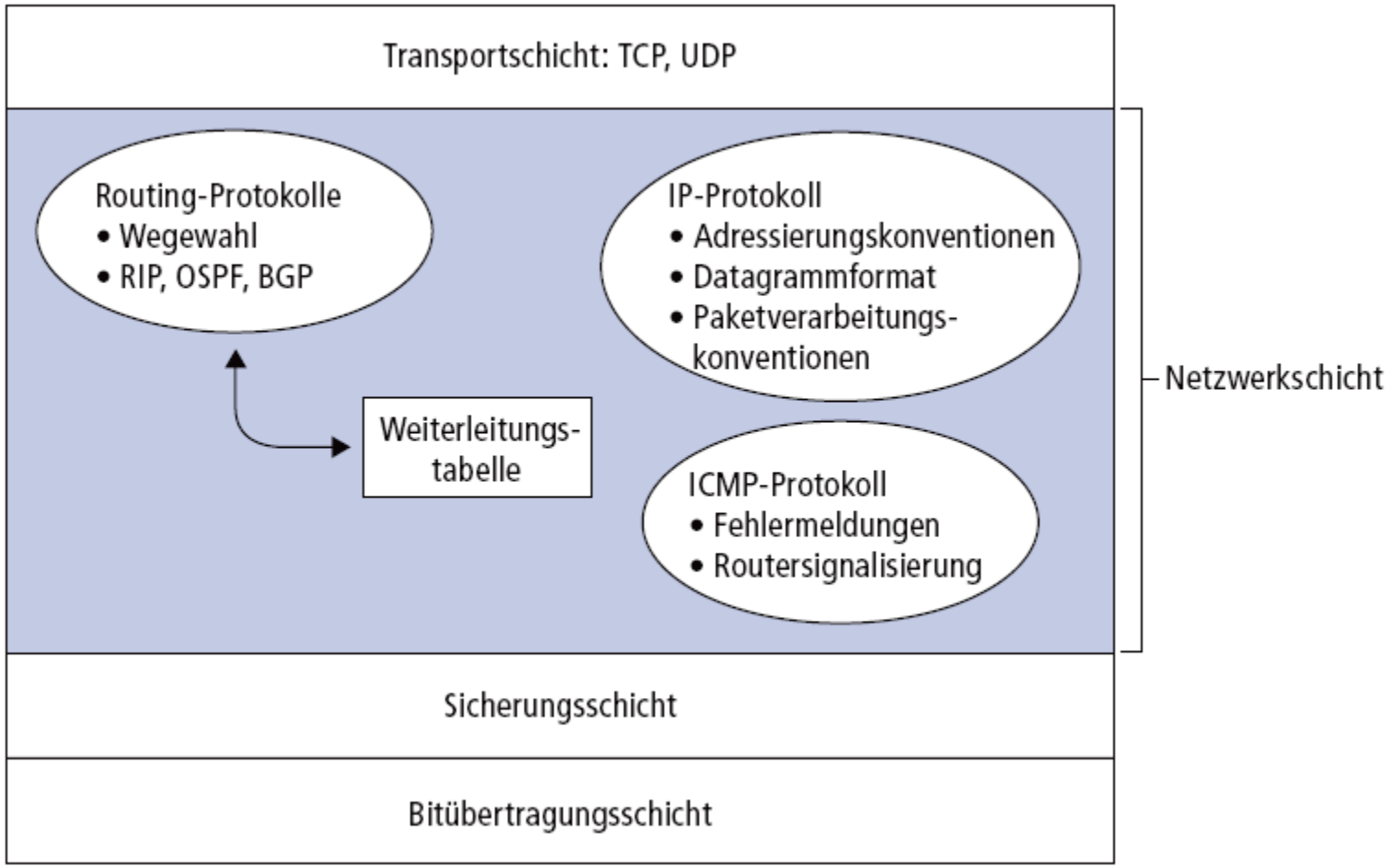
4.3.4 Puffern im Inputport

- Wenn die Switching Fabric ein Paket nicht direkt weiterleiten kann, muss dieses im Eingangsport gepuffert werden
- Dort kann es ein Paket blockieren, welches eigentlich bereits durch die Switching Fabric geleitet werden könnte
  → **Head-of-Line (HOL) Blocking**

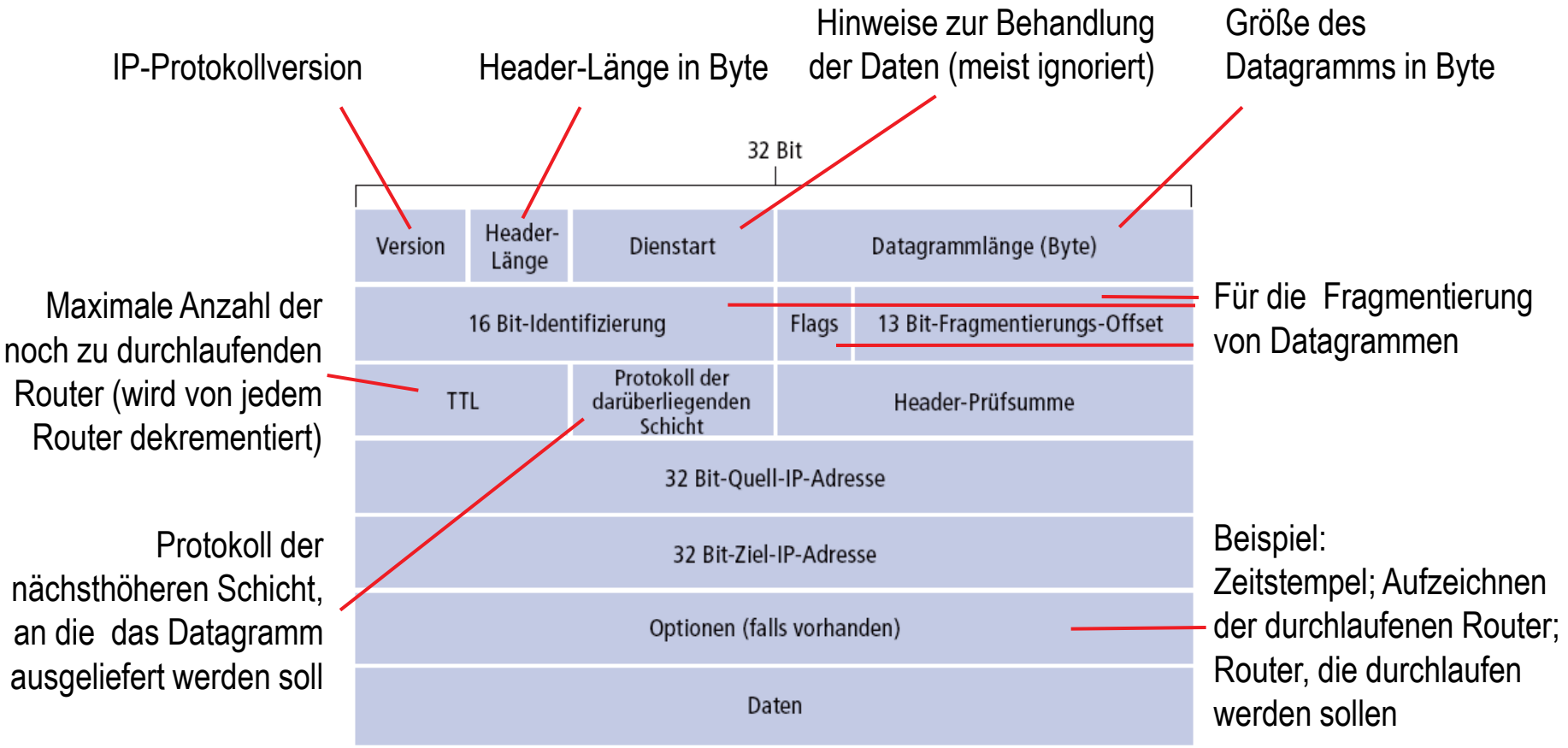


4.4 IP: Internetprotokoll

Die Netzwerkschicht des Internets:



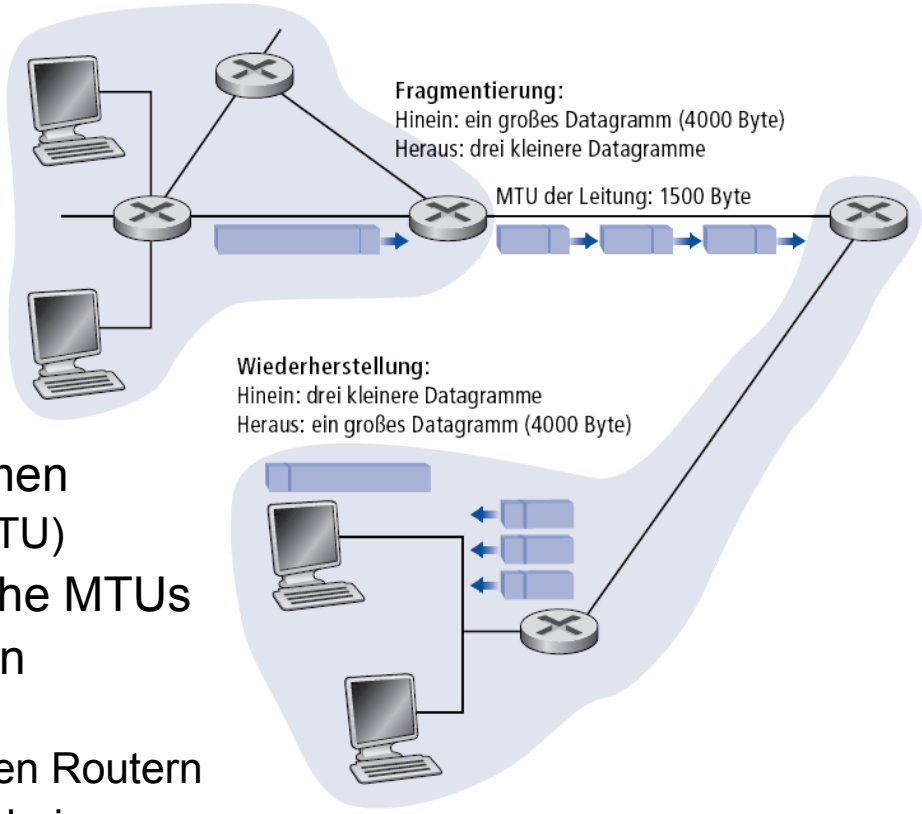
4.4.1 IP-Datagrammformat



Wie viel Overhead entsteht bei Verwendung von TCP?

→ 20 Byte für den TCP-Header, 20 Byte für den IP-Header= 40 Byte + Overhead auf der Anwendungsschicht

4.4.1 IP-Datagramm-Fragmentierung



- Links haben eine Maximalgröße für Rahmen
  - Genannt Maximum Transmission Unit (MTU)
- Verschiedene Links haben unterschiedliche MTUs
- IP-Datagramme müssen unter Umständen aufgeteilt werden
  - Aufteilung (Fragmentierung) erfolgt in den Routern
  - Zusammensetzen (Reassembly) erfolgt beim Empfänger
  - IP-Header enthält die notwendigen Informationen hierzu

4.4.1 IP-Datagramm-Fragmentierung

Beispiel: IP-Datagramm mit 4000 Byte (inklusive 20 Byte IP-Header), MTU des nächsten Links = 1500 Byte

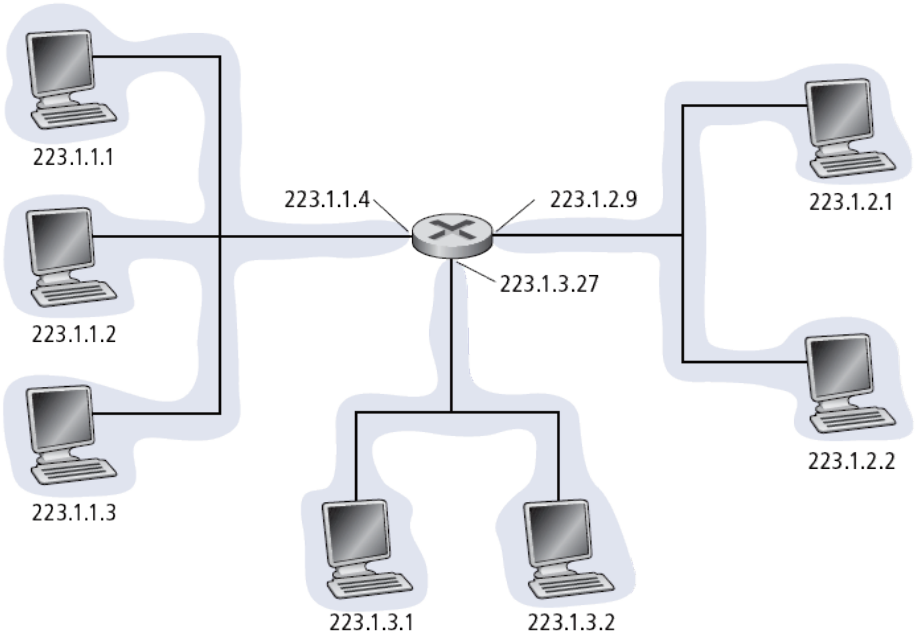
Fragment	Bytes	ID	Offset	Flag
1. Fragment	1.480 Byte im Datenfeld des IP-Datagramms	Identifizierung = 777	Offset = 0 (d.h., die Daten sollten beginnend bei Byte 0 eingefügt werden)	Flag = 1 (d.h., da kommt noch mehr)
2. Fragment	1.480 Datenbytes	Identifizierung = 777	Offset = 185 (d.h., die Daten sollten bei Byte 1.480 beginnend eingefügt werden; beachten Sie, dass 185 · 8 = 1.480)	Flag = 1 (d.h., da kommt noch mehr)
3. Fragment	1.020 Datenbytes (= 3.980 – 1.480 – 1.480)	Identifizierung = 777	Offset = 370 (d.h., die Daten sollten beginnend bei Byte 2.960 eingefügt werden; beachten Sie, dass 370 · 8 = 2.960)	Flag = 0 (d.h., es ist das letzte Fragment)

4.4.1 IP-Datagramm-Fragmentierung

- Praktisch:
  - Endsystem/Anwendung muss sich keine Gedanken über die Größe von MTUs verschiedener Links auf dem Weg vom Sender zum Empfänger machen
- Entspricht dem Prinzip einer geschichteten Architektur
- Aber:
  - Aufwand in den Routern
  - Wenn ein Fragment verloren geht, ist das ganze Datagramm verloren
- Daher: *Fragmentation considered harmful!*
- Lösung: Bestimmen der kleinsten MTU des Weges (Path MTU)
  - Setze **DF (Don't-Fragment-Bit)** im Header des IP-Paketes
  - Wenn fragmentiert werden soll, wird das Paket verworfen und der Sender per ICMP benachrichtigt
  - Sender wählt dann kleinere MTU
  - Wiederholen, bis akzeptable MTU gefunden wurde

4.4.2 IPv4-Adressierung

- IP-Adresse: 32-Bit-Kennung für das Interface (Schnittstelle) eines Endsystems oder eines Routers
- Interface: Verbindung zwischen dem System und dem Link
  - Wird normalerweise durch eine Netzwerkkarte bereitgestellt
  - Router haben typischerweise mehrere Interfaces
  - Endsysteme können ebenfalls mehrere Interfaces haben
  - Jedes Interface besitzt eine IP-Adresse

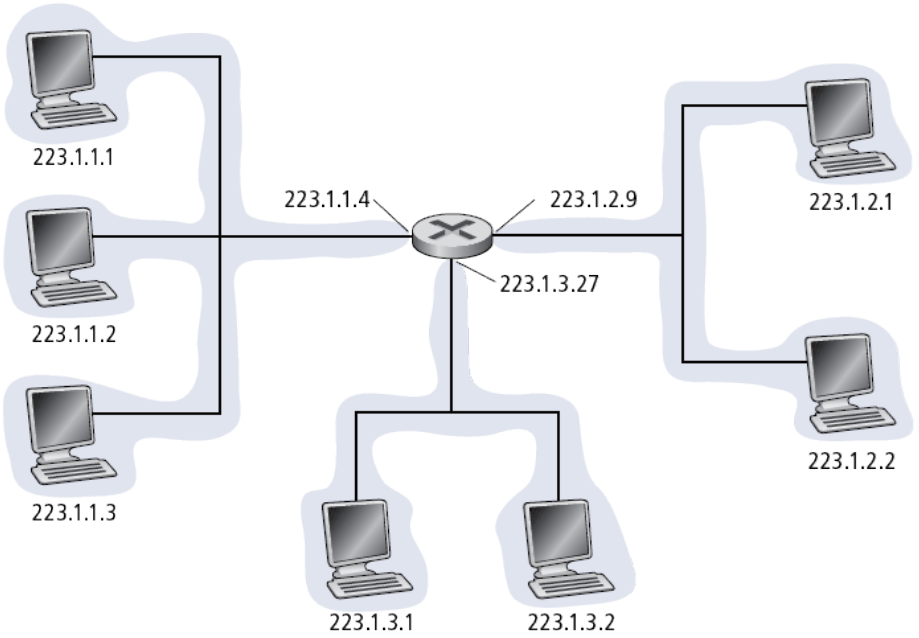


223.1.1.1 = 11011111.00000001.00000001.00000001

223 1 1 1

4.4.2 IPv4-Adressierung

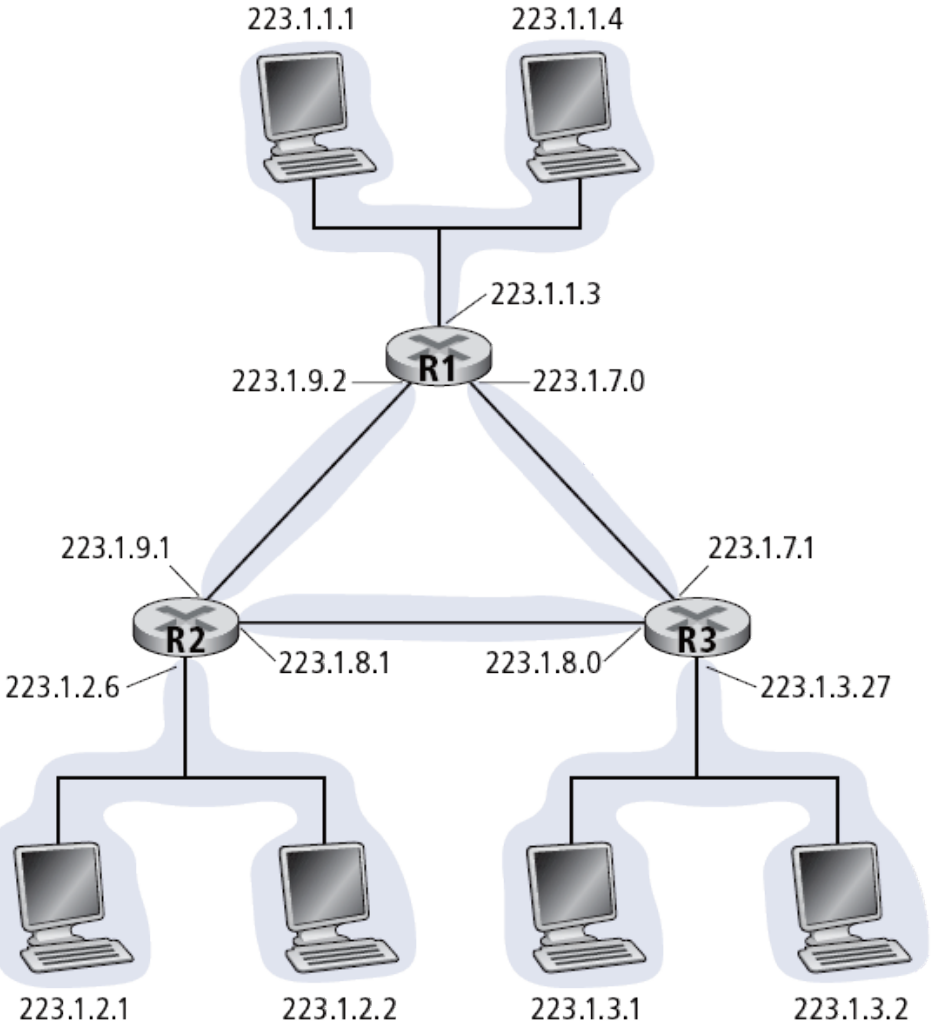
- IP-Adresse:
  - Zwei Bestandteile:
    - **netid**: die oberen Bits der Adresse, identifiziert ein Netzwerk
    - **hostid**: die unteren Bits der Adresse, identifiziert ein Interface eines Systems
- Was ist ein (Sub-)Netzwerk?
  - Alle Interfaces mit derselben netid formen ein Netzwerk
  - Alle Interfaces eines Netzwerkes können sich direkt (ohne einen Router zu durchqueren) erreichen



Drei IP-Netzwerke, die mit einem Router verbunden sind. Die netid steht hier in den oberen 24 Bit.

Subnetzmaske: `/24` oder `255.255.255.0`

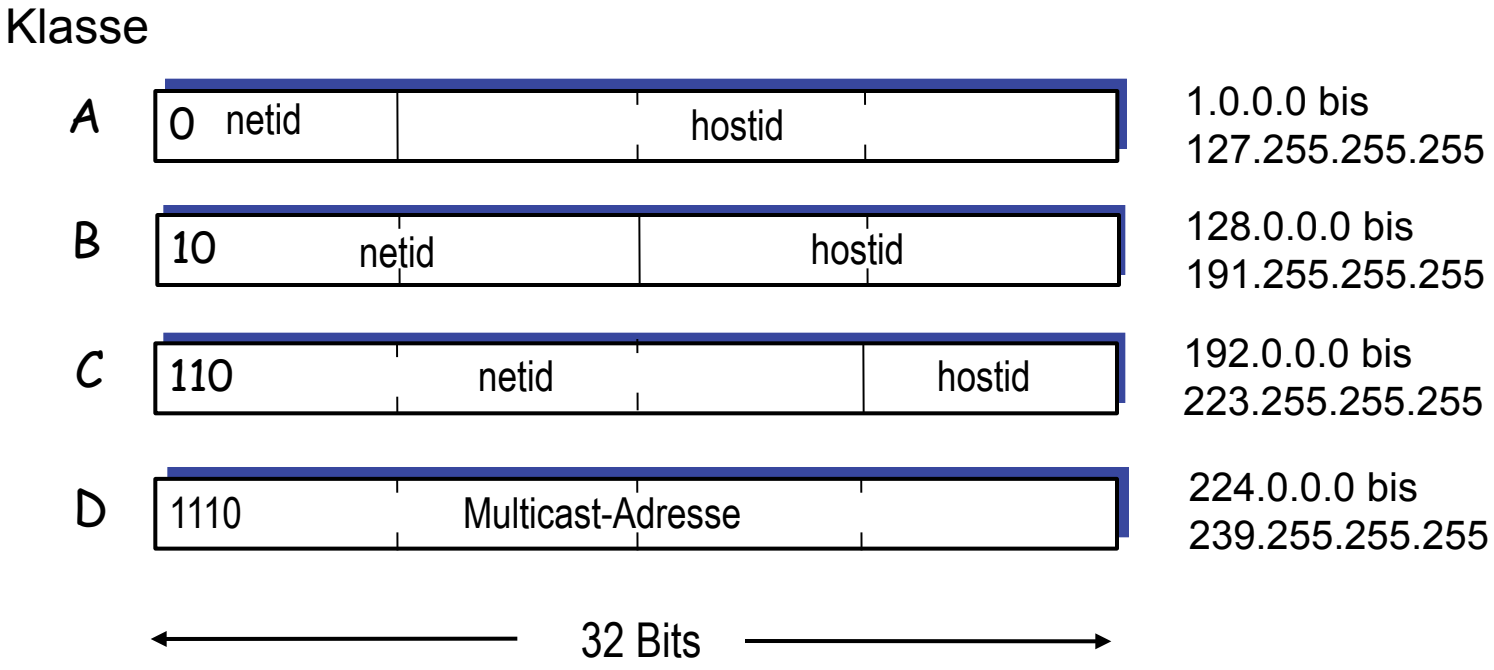
4.4.2 IPv4-Adressierung



Drei Router verbinden sechs /24-Subnetzwerke untereinander.

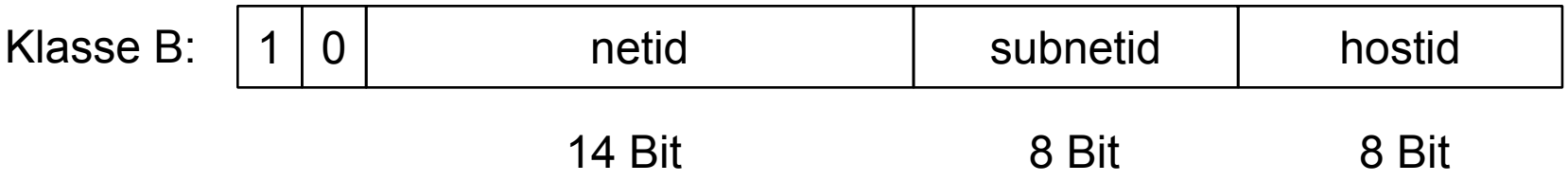
4.4.2 IP Adressierung - Adressklassen

- Früher wurden IP-Adressen in Adressklassen aufgeteilt
- Die Klasse bestimmte das Verhältnis der Längen netid/hostid
- Dies nennt man „classfull“ addressing (klassenbasierte Adressierung)



4.4.2 Adressierung von Subnetzen

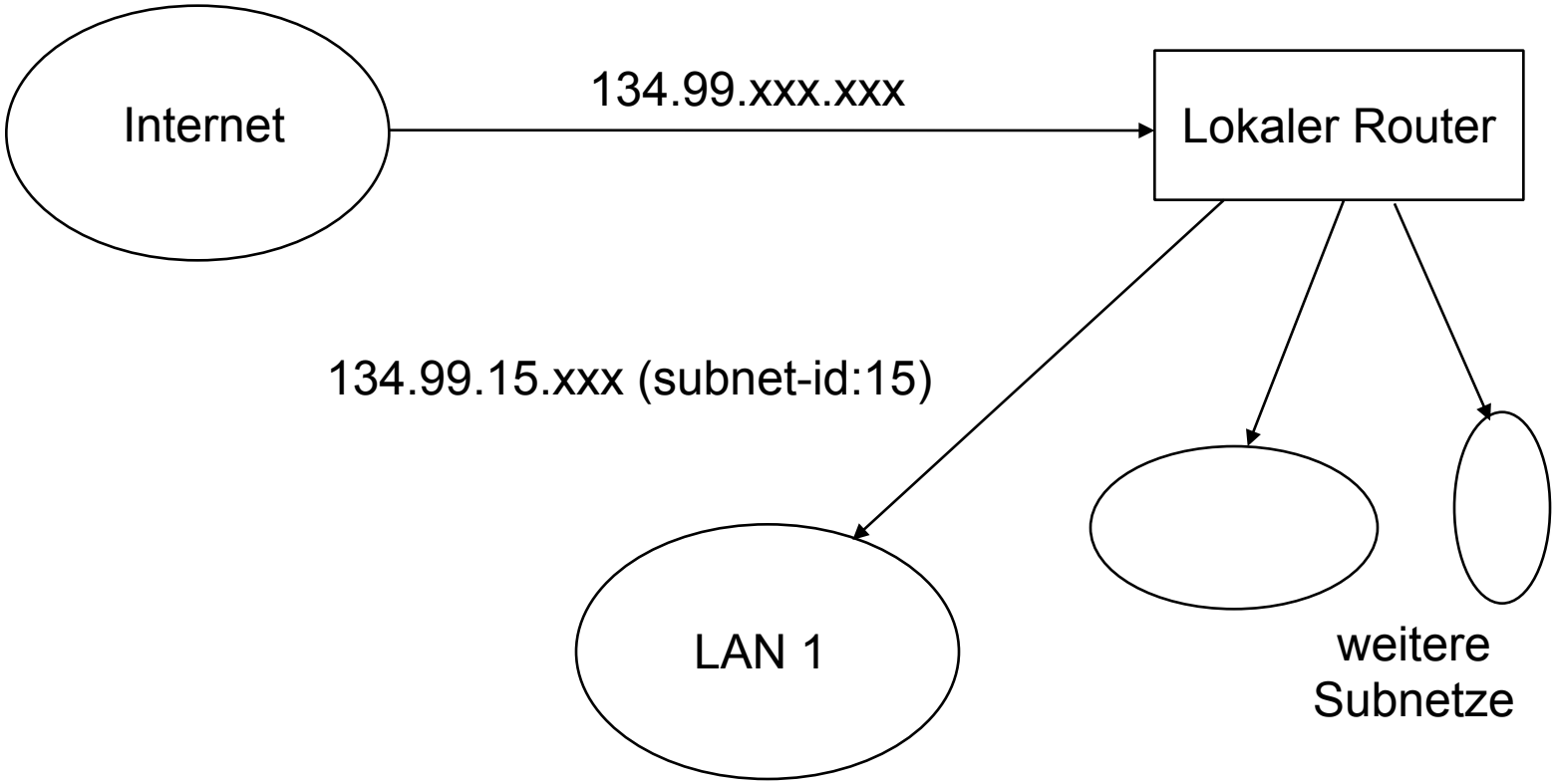
- Klasse-A- und –B-Adressen haben Platz für mehr Endsysteme, als man in einem Netzwerk sinnvoll unterbringen kann
- Daher teilt man die hostid weiter auf, z.B. so:



- Die Unterteilung (subnetid, hostid) ist eine lokale Entscheidung und wird von der Organisation vorgenommen, der die netid zugeordnet wurde

4.4.2 Adressierung von Subnetzen

- Die subnetid ist außerhalb des Netzwerkes, für das sie verwendet wird, nicht sichtbar:



4.4.2 Adressierung von Subnetzen

- Subnetzmaske (subnet mask)
  - Wird für jede IP-Adresse eines Systems im System gespeichert
  - Sie identifiziert, welcher Teil der Adresse zur subnetid und welcher zur hostid gehört

	16 Bit	8 Bit	8 Bit
Beispiel	1111111111111111	11111111	00000000

subnet mask: 0xffffff00=255.255.255.0
oder auch /24

- Die eigene IP-Adresse in Verbindung mit der Subnetzmaske erlaubt Rückschlüsse darüber, wo sich eine andere IP-Adresse befindet:
  - im selben Subnetz (also direkt erreichbar)
  - im selben Netzwerk, aber in einem anderen Subnetz
  - in einem anderen Netzwerk

4.4.2 Beispiel für die Verwendung von Subnetzmasken

- Gegeben:
  - Eigene IP-Adresse: `134.155.48.10`
  - Subnetzmaske: `255.255.255.0`
  - Adresse A: `134.155.48.96`, Adresse B: `134.155.55.96`

- Überprüfen der beiden Adressen:
  - `134.155.48.10` & `255.255.255.0` = `134.155.48.0`
  - `134.155.48.96` & `255.255.255.0` = `134.155.48.0` → identisch, gleiches Subnetz
  - `134.155.55.96` & `255.255.255.0` = `134.155.55.0` → verschieden, anderes Subnetz