

Network Layer - Forwarding, IPv4

① Forwarding und Routing

Network Layer:

- Ende-zu-Ende Beziehung zwischen Sender und Empfänger
- Sender verpackt ein Transport Layer Segment in Datagramm
- Empfänger liefert das Datagramm an Transport Layer aus
- Router interessiert sich nicht für Schicht 4/5 und kümmert sich nur um Weiterleitung zum Zielhost

Aufgaben von Network Layer

- Adressierung von IP-Adressen
- Forwarding:
 - ⇒ zu welchem Ausgangsinterface muss ein Datagramm in Richtung Ziel weitergeleitet werden
 - ⇒ bei Router oft in HW implementiert wird
- Routing (oft : Wegewahl):
 - ⇒ berechnung Wege in Netz
 - ⇒ Eintragen der Ergebnisse in Weiterleitungstabelle
 - ⇒ Routing protokolle
 - ⇒ meist in SW implementiert
- IP ist verbindungslos

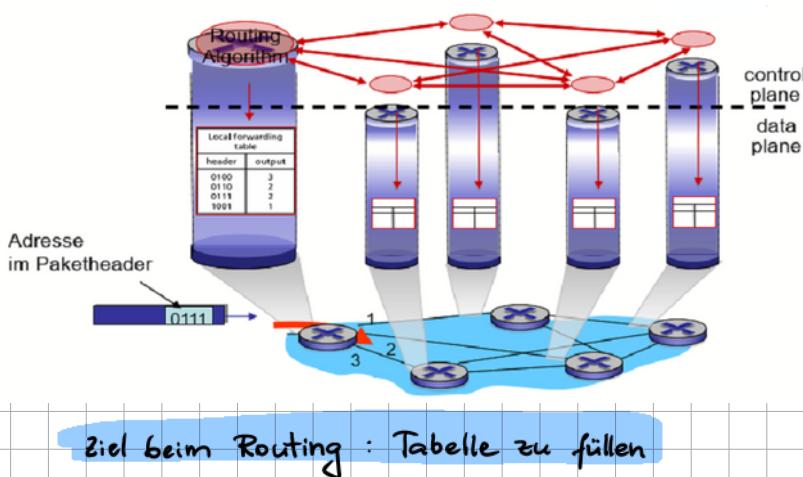
Analogie:

Routing

Navigationsystem
berechnet die Reiseroute

Forwarding

Navigation teilt Fahrer an eine Kreuzung
mit, ob er links oder rechts abbiegen muss



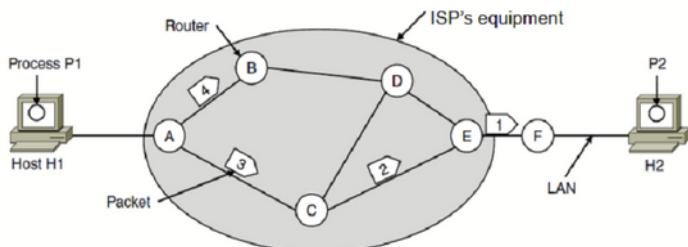
Routing (Control plane)

- automatische Wegeberechnung: Netzwerkweite Funktion
- Routingnachrichten zwischen Routern

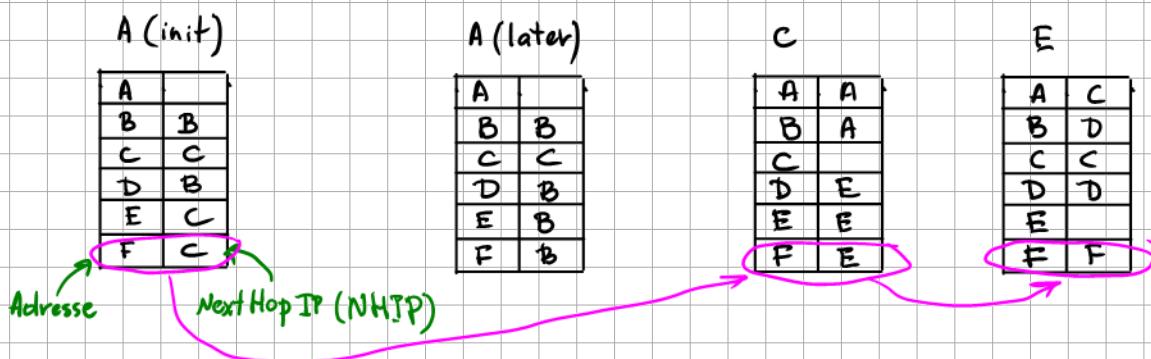
Forwarding (Data plane)

- Weiterleitung: lokale Funktion jedes Routers

JP: verbindungsloses Forwarding



- Weiterleitung des Pakets nur anhand der Ziel-IP
- Bei Änderung der Tabelle können Pakete verschiedene Pfade folgen



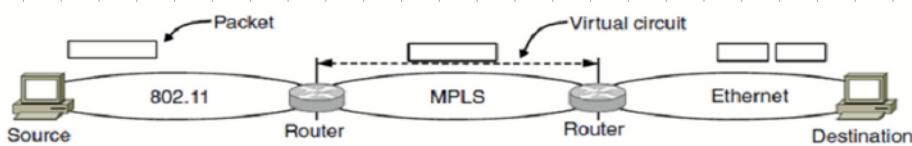
Zusammenarbeit von heterogenen Netzen

Ein heterogenes Netzwerk ist dadurch charakterisiert, dass es auf einer oder mehreren Schichten Heterogenität aufweist

Vorteile:

- vorhandene Hard- & Software kann weiter benutzt und neue Produkte direkt in den Netzwerk eingebunden werden

- IP ist Bindeglied (cbz. zBeno)
- die Link-Layer kann unterschiedlich sein



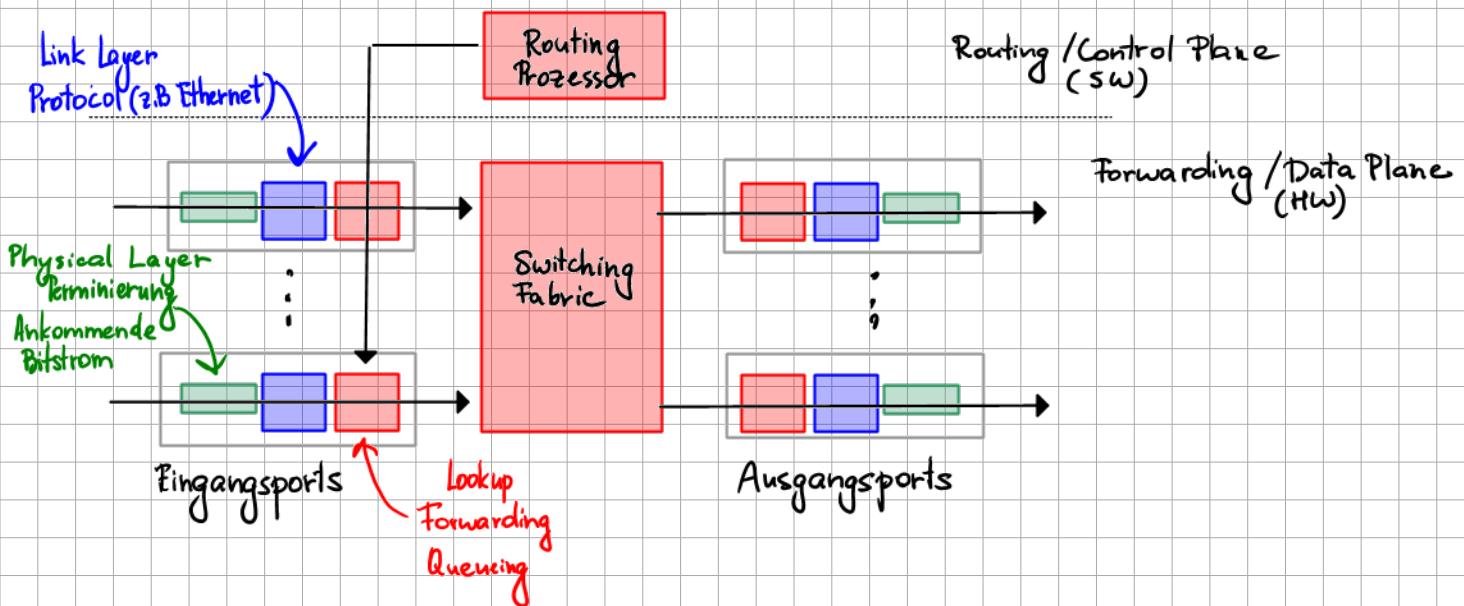
Best Effort

- keine Bevorzugung bestimmte Pakete
- Jeder Router tut "sein Bestes"
- Reihenfolge, Bandbreite usw nicht garantiert
- Router speicher keine Info über aktuell laufende Ende-zu-Ende Beziehung

(2) Funktionsweise eines Routers

Architektur eines Routers

- Eingangsport: Pufferung ankommender IP-Pakete, terminiert Link Layer
- Fabrik: "Netzwerk innerhalb Router", Weiterleitung zu passendem Ausgangsport
- Ausgangsports: Puffern bis Übertragung möglich, Link / PHY Funktionalität
- Routingprozessor: Ausführung der Routing protokolle



- Dezentral: Nachschlagen des Ziels im Eingangsport
- Verarbeitung mit "Line Speed" == Übertragungskapazität des ankommenden Links
- Queueing: falls Ankunfrate schneller als Weiterleitung durch Fabric

Bei IP wird Ausgangsport nur anhand der IP-Zieladresse bestimmt!

Forwarding Table

Die lokale Forwarding Table ist in einem Tabelleingetragen:

Lokale Forwarding Table	
Header	Interface
Adresse-range 1	3
Adresse-range 2	2
Adresse-range 3	2
Adresse-range 4	1

Problem: Viele Einträge
da IP 32 Bit hat

Lösung: Aggregation / Hierarchie
In die Forwarding Tabelle werden Adressbereiche eingetragen

Weiterleitung nach Zieladresse

- Router leitet nach Bereichen weiter:
- ⊕ Skalierbarkeit, da nicht jede einzelne Adresse Tabellenplatz belegt

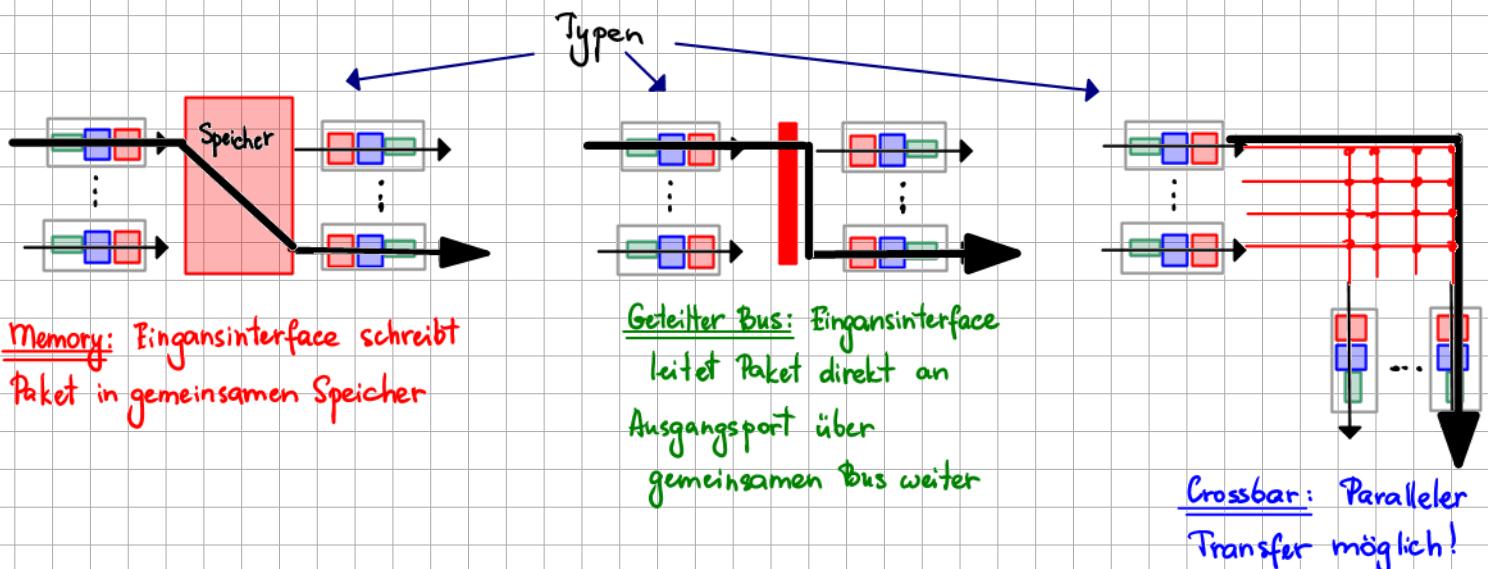
Zieladresse (IP)		Hex	Ausgangsport
1100 1000 00010111 00010000 00000000 bis		200. 23. 16. 0	0
1100 1000 00010111 00010111 11111111		200. 23. 23. 255	
1100 1000 00010111 00011000 00000000 bis		200. 23. 24. 0	1
1100 1000 00010111 00011000 11111111		200. 23. 24. 255	
1100 1000 00010111 00011001 00000000 bis		200. 23. 25. 0	2
1100 1000 00010111 00011111 11111111		200. 23. 47. 255	
sonst			3

Largest Prefix Matching

- Anstatt Adressbereiche anzugeben, verwendet man Präfixe
- Largest Prefix Matching:
 - ⇒ Nachschlagen einer Ziel IP (32 Bit) in Forwardingtabelle des Routers
 - ⇒ Suche längsten Adresspräfix, der mit der Zieladresse übereinstimmt

Switching Fabric

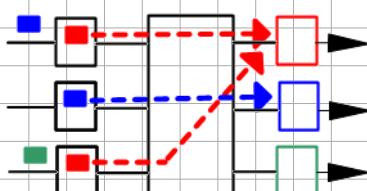
- Aufgabe:
- Weiterleitung von Paketen vom Eingangs - zum Ausgangsport
 - Ideal: Durchsatzrate = Summe der Ankunftsrraten aller Eingangsports



Queueing am Eingangsport

- Nötig falls Fabric langsamer als Ankunftsrate
- Problem: Head-of-the-Line Blocking (vorderes Paket in Queue blockiert andere Pakete)

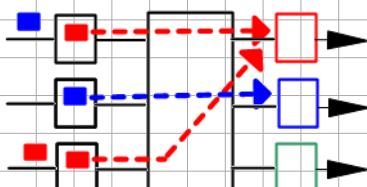
Beispiel:



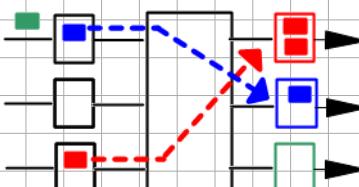
- 2 rote Pakete müssen gleichzeitig zum selben Ausgangsport
- Annahme: Unteres rotes Paket muss warten
- Aber grünes muss auch dann warten, obwohl sein Ausgangsport frei ist

am Ausgangsports

- Nötig falls Ankunftsrate von Switch Fabric die Übergangsrate des Ausgangs links übersteigt
- Ursache für:
 - ✓ Delay / Verzögerung
 - ✓ Paketverluste



Mehrere Pakete müssen zum gleichen Ausgang



ein Paket später

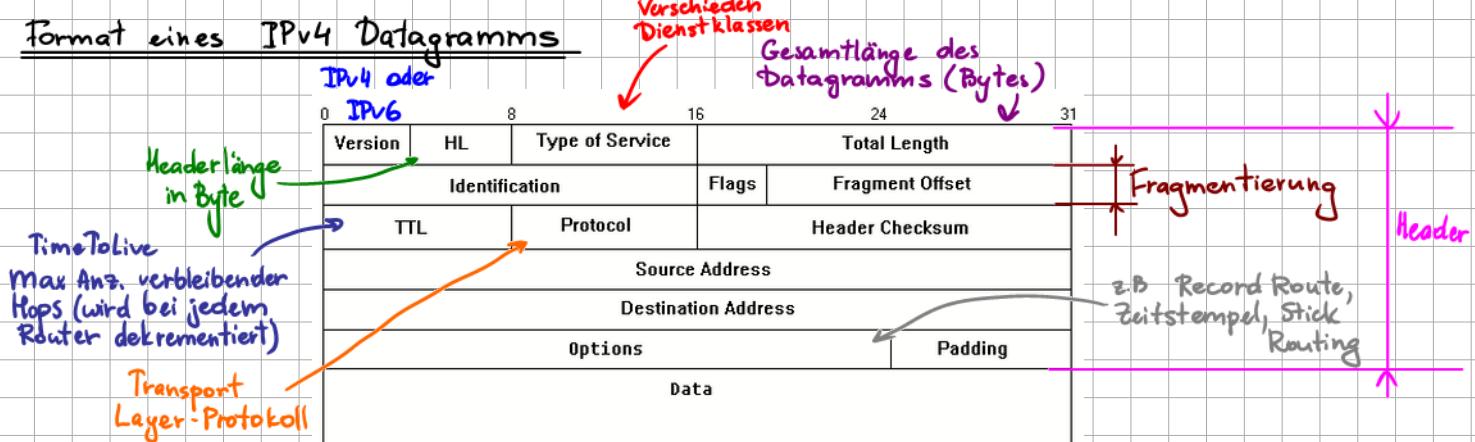
Paket hat keinen Platz mehr

→ Paketverlust

(3) Internet Protocol IPv4

- IP:
- addressing conversions
 - datagram format
 - packet handling conversions

Format eines IPv4 Datagramms



Wie viele Overheads? 20 Bytes für IP-Header (unnötige Daten)

IP Fragmentierung

- **Max Transfer Unit (MTU)**: Verschiedene Link Layer Technologien erlauben verschiedene max. Paketgröße
- IP-Datagramm > MTU:
 - ⇒ Router / Host zerlegt in kleinere "Fragmente"
 - ⇒ Zusammenbau am End-Host, nicht im Netz!
 - ⇒ IP Header Bits um Fragmente zu identifizieren und wieder zusammenzufügen



IP-Fragmentierung : Wiederzusammenfügung

- **16 Bit Identifier**: identisch für alle Fragmente eines Pakets
- **Fragmentation Flag**: 0 markiert das letzte Fragment eines Pakets
- **Offset**: Byteposition innerhalb des Pakets, an die das Fragment gehört

Bsp:

⇒ 4000 Byte Datagramm, wo Header-Länge = 20 Byte

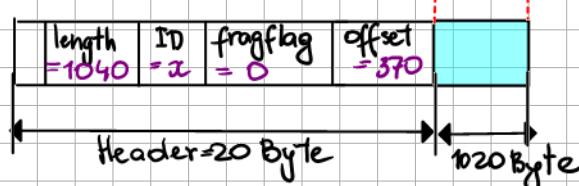
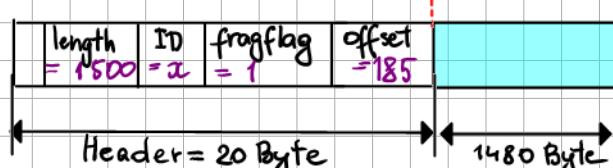
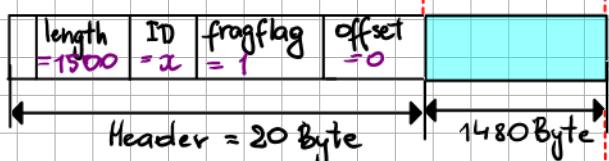
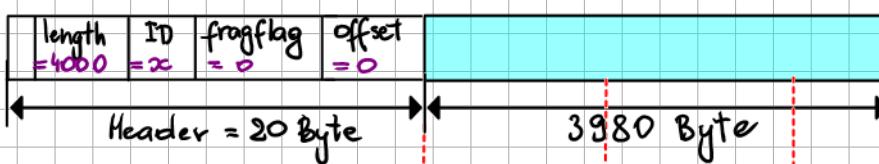
⇒ MTU = 1500 Byte

Das Datagramm muss durch 8 dividiertbar sein =>

$$1500 - 20 = 1480 \text{ Header}$$

$$1480 \% 8 = 0$$

$$\frac{1480}{8} = 185$$



Datagramm = 4000 Byte

Data = $4000 - 20 = 3980$ Byte
Header

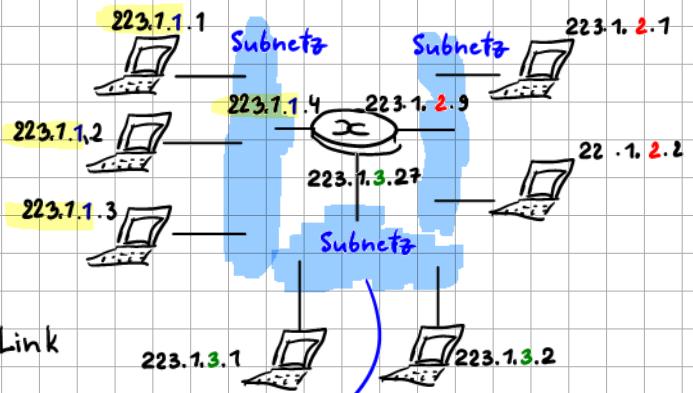
$$\frac{3980}{1480} = 2.69$$

MTU-Header

$$\begin{aligned} 1480 + 20 &= (\text{Paket 1}) \\ 1480 + 20 &= (\text{Paket 2}) \\ 3980 - 1480 \cdot 2 + 20 &= 1040 \\ &(\text{Paket 3}) \end{aligned}$$

IP-Adressierung

- IP-Adresse :
- 32 Bit
 - identifiziert Host im Internet
 - gehört zu Interface



- Interface:
- Verbindung zwischen Host / Router und Link
 - Router haben mehrere Interfaces
 - jeder Interface benötigt eine IP-Adresse

- IP Subnetz:
- Hosts teilen sich gleichen IP Adresspräfix
 - Hosts können sich ohne Router erreichen
 - Bsp: Ethernet, WLAN ...

Adresse eines IP-Subnetzes:

- Subnetz ist über gemeinsamen Präfix adressierbar!
- Subnetzmaske (gelb) - Länge des gemeinsamen Präfix
- Hostanteil : Bits der IP-Adresse, die sich für jeden Host unterscheiden

Bsp : 223.1.3.0/24 Die ersten 24 (8x3) Bits sind für alle Hosts des Subnetzes gleich

1101 1111. 0000 0001. 0000 0011. 223 . 1 . 3 . 0/24

⊕ man muss nur Subnetzadressen in den Routingtabellen haben

Classfull Adressierung (Netzklassen)

- Adressbereiche werden an Firmen, Uni usw. übergeben

	31	24	16	8	0
Class A	0	Net-ID		Host-ID	
Class B	1 0			Host-ID	
Class C	1 1 0		Net-ID		Host-ID
Class D	1 1 1 0			multicast address	
Class E	1 1 1 1 0				reserved for future use

	Klasse A - Netz	Klasse B - Netz	Klasse C - Netz
Netz-ID	8 Bit = 1 Byte	16 Bit = 2 Byte	24 Bit = 3 Byte
Host-ID	24 Bit = 3 Byte	16 Bit = 2 Byte	8 Bit = 1 Byte
Netzmaske	255.0.0.0	255.255.0.0	255.255.255.0
Adressklassen-ID (= Feste Bits im 1. Byte, 1. Quad)	0	10	110
Wertebereich (theoretisch)	0.0.0 bis 127.255.255.255	128.0.0 bis 191.255.255.255	192.0.0 bis 223.255.255.255
Anzahl der Netze	$128 (= 2^7)$	$16384 (= 2^{6*256})$ $= 64*256$	$2097152 (= 2^{5*256*256})$ $= 32*256*256$
Anzahl der Rechner im Netz	$16777216 (= 256^3)$	$65536 (= 256^2)$	$256 (= 256^1)$

	Klasse D	Klasse E
Adressklassen-ID	4 Bit = "1110"	5 Bit = "11110"
keine Netz-ID, sondern:	28 Bit-Identifikator	27 Bit-Identifikator
Wertebereich	224.0.0 bis 239.255.255.255	240.0.0 bis 247.255.255.255
Anwendungen	für Multicast-Gruppen	reservierte Adressen für Zukünftiges

Grundsätzlich gilt:

- alle Rechner mit der gleichen Netzadresse gehören zu einem Netz und sind untereinander erreichbar
- zur Koppelung von Netzen unterschiedlicher Adresse wird spezielle HW- oder SW-Komponente, ein Router, benötigt
- je nach Zahl der koppelnden Rechner wird die Klasse ausgewählt

Classes Interdomain Routing (CIDR) : Subnetzteil einer IP-Adresse kann beliebige Länge haben

- Notationen : Präfixnotation : 200.23.16.0/24

Mit Netzmase : Adresse: 200.23.16.0 + Netzmase 255.255.255.0
(Netzmase gibt an, welche Bits zum Subnetz gehören)

Spezielle IPv4 Adressen

- localhost eigene PCs : 127.0.0.1

- Private IPv4 Adressen :

→ global nicht sichtbar, nur lokal im eigenen administr. Netz zu verwenden

⇒ 10.0.0.0/8

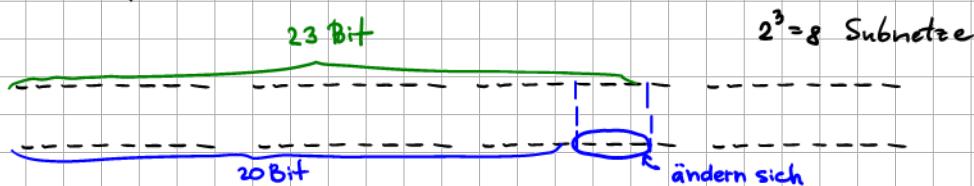
⇒ 172.16.0.0/12

⇒ 192.168.0.0/16

Bsp /20 wird auf /23 Netze unterteilt

ISP's block	20 Bit			23 Bit	200.23.16.0
	1100 1000	0001 0111	0001 0000		
Organisation 0 :	1100 1000	0001 0111	0001 0000	0000 0000	200.23.16.0
Organisation 1 :	1100 1000	0001 0111	0001 0010	0000 0000	200.23.18.0
Organisation 2 :	1100 1000	0001 0111	0001 0100	0000 0000	200.23.20.0
...					
Organisation 7 :	1100 1000	0001 0111	0001 1110	0000 0000	200.23.30.0

Ü /23 auf /20 : Wie viele Subnetze?



Aufteilung in Subnetze

Netzwerkanteil in Bit	Hostanteil in Bit	Subnetzanzahl *)	Hostanzahl **)	Subnetzmaske
8	24	1	16777216	255.0.0.0 Klasse A
9	23	2	128*65536	255.128.0.0
10	22	4	64*65536	255.192.0.0
11	21	8	32*65536	255.224.0.0
12	20	16	16*65536	255.240.0.0
13	19	32	8*65536	255.248.0.0
14	18	64	4*65536	255.252.0.0
15	17	128	2*65536	255.254.0.0
16	16	1	65536	255.255.0.0 Klasse B
17	15	2	128*256	255.255.128.0
18	14	4	64*256	255.255.192.0
19	13	8	32*256	255.255.224.0
20	12	16	16*256	255.255.240.0
21	11	32	8*256	255.255.248.0
22	10	64	4*256	255.255.252.0
23	9	128	2*256	255.255.254.0
24	8	1	256	255.255.255.0 Klasse C
25	7	2	128	255.255.255.128
26	6	4	64	255.255.255.192
27	5	8	32	255.255.255.224
28	4	16	16	255.255.255.240
29	3	32	8	255.255.255.248
30	2	64	4	255.255.255.252

Wie bekommt man ein ISP einen IPv4 Adressbereich?

Registry ICANN (Internet Corporation for Assigned Names and Numbers)

Zuständig für:

- Vergabe von IPs
- Domain Name System (DNS), Root Domains

(4) ARP, ICMP, DHCP

Wie weist man Hosts eine IP zu?

⇒ Manuell (durch SS)

⇒ Automatisch durch DHCP (Dynamic Host Configuration Protokoll):

- Plug-and-Play
- IP wird automatisch durch Server zugewiesen

DHCP Server leitet IP an Host aus Pool von Adressen aus. Host kann zugewiesene IP ggf. verlängern. Schritt 4

DHCP Funktionsweise:

- Host sucht einen DHCP Server : **DHCP Discover (optional)**
Ziel IP: 255.255.255.255
- DHCP Server antwortet mit **DHCP Offer (optional)** :
Ziel IP: 255.255.255.255
- Host fordert explizit IP an : **DHCP Request**
- DHCP Server weist Adresse zu : **DHCP ACK**

Per DHCP lassen sich weitere Parameter konfigurieren:
DNS Server, Subnetzmaske,
IP des Gateways

ICMP (Internet Control Message Protokoll) :

- error reporting
- router "signalling"
- Austausch von Info zwischen Host und Routern :
 - bei Fehler sendet Router einen Fehlerbericht, z.B. "unreachable Host, Port, Protokoll"
 - Echo Request / Reply : Ping
- ICMP wird als IP-Paket versendet
- ICMP Nachricht enthält Type Code + die ersten 8 Bit des IP-Pakets, das den Fehler verursacht

ARP (Adress Resolution Protocol)

- 32-Bit IP : \Rightarrow Network Layer Adresse für ein Interface
 - Forwarding auf Schritt 3
- 48-Bit MAC : \Rightarrow fest verbunden mit Netzwerkadapter
 - Verwendung: lokal, für Zustellung auf einem "Link"

Aufgaben von ARP:

\Rightarrow auf dem Weg zum Ziel muss IP-Paket über mehrere Links weitergeleitet werden

\rightarrow jeder Router / Host schlägt Ausgangsport nach und leitet dann Paket weiter

\Rightarrow Aber welche Ziel-MAC-Adresse gehört zum Next-Hop Router/ Host ?

\Rightarrow Nötig: Übersetzen von IP in MAC

ARP Auflösung: IP zu MAC:

- IP Knoten = Hosts / Router // nicht Switch!
- jede IP Knothe verwaltet eine ARP Tabelle ($\langle \text{IP}, \langle \text{MAC} \rangle, \langle \text{TTL} \rangle \rangle$)
- TTL (Time to Live) - Zeit nachdem Eintrag ungültig wird (oft nach 20 min)