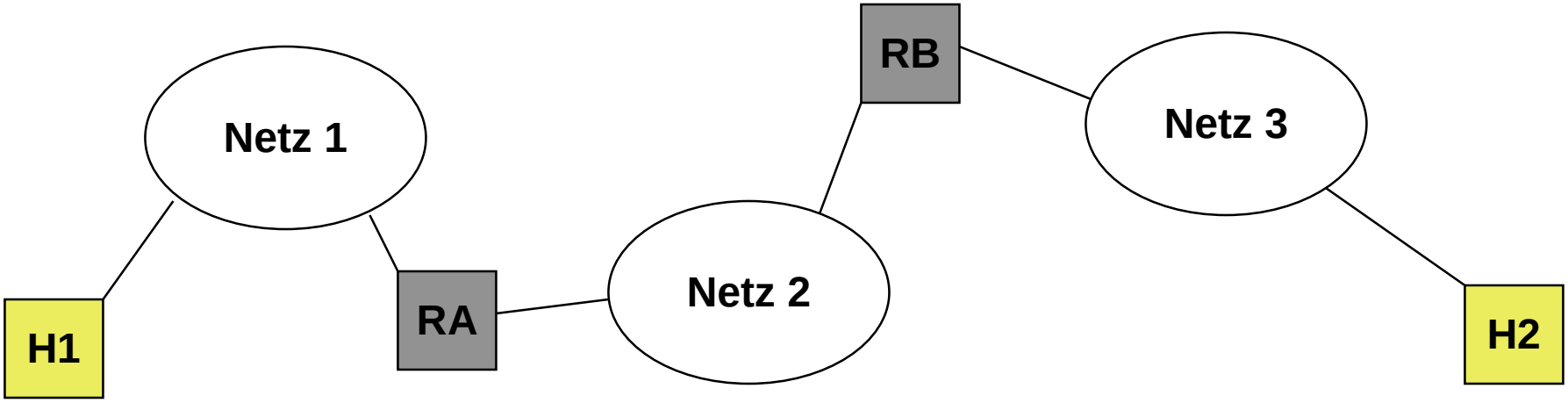
Das Internet Protocol (IP): Funktionen, Datenformate und Adressierung

Konzept und Architekturmodell des Internet

- Das Internet beruht auf dem Zusammenschluss von teilautonomen Subnetzen mittels Verbindungsrechnern (Routern).
- Router leiten den Datenverkehr gemäss einer Netzwerkadresse, nicht einer Endsystemadresse.
- Diese Form der Datenweiterleitung ist transparent für Benutzer.



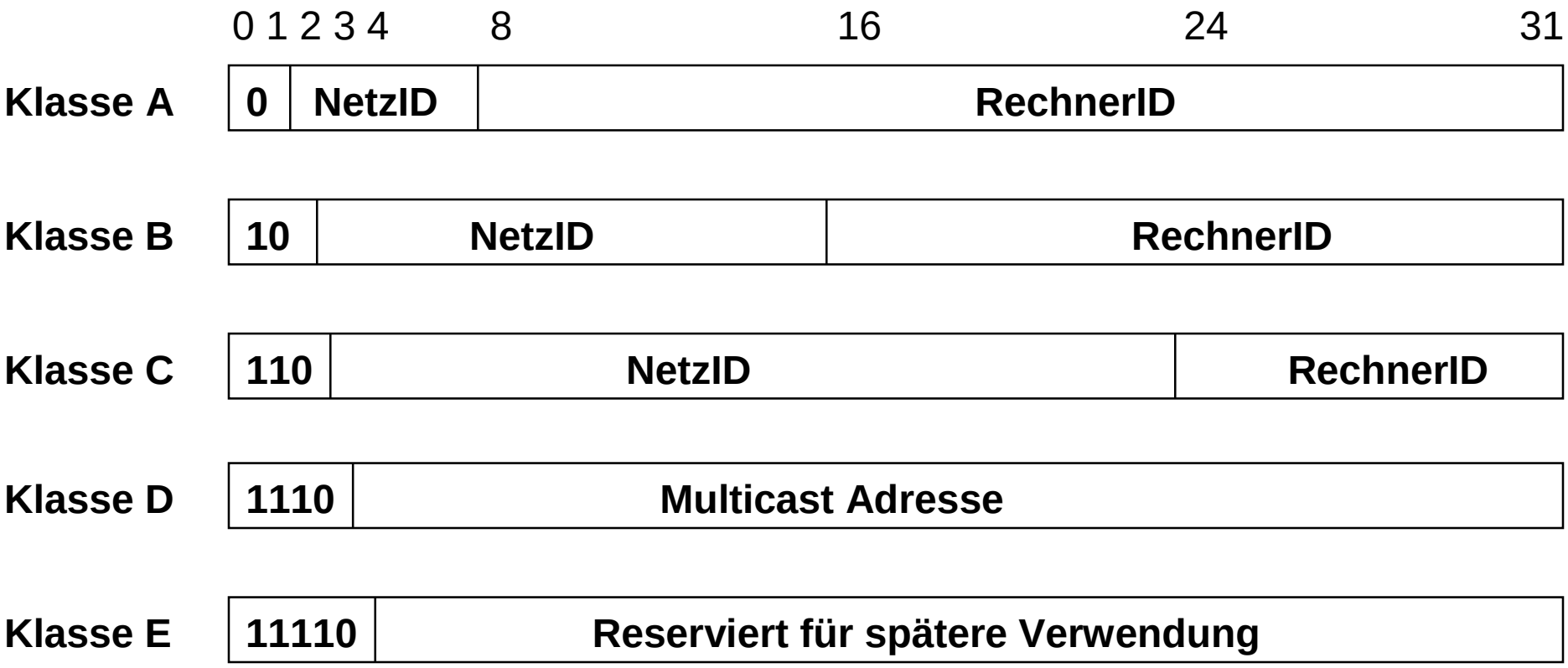
Adressierung im Internet: Namen, Adressen, Routen

- Namen: *Identifikation* bezüglich eines Kontext, z.B. "Zürich"
- Adressen: *Ort* bezüglich eines Kontext, z.B. "PLZ 8000"
- Routen: *Weg* zum adressierten Ort, z.B. Landkarte

- Im Internet werden von Benutzern Rechner oder Rechnernetze benannt (DNS) oder adressiert (IP-Adressen).

- Alle Internetadressen sind 32-Bit Werte, die in eine Netznummer und eine Rechnernummer bezüglich dieses Netzes unterteilt sind.

Adressierung im Internet: Adresstypen



Adressierung im Internet: Darstellung

- Darstellung als 4 Oktette in Dezimalnotation, getrennt durch einen Punkt, z.B. 129.132.66.1

- Klasse A zwischen 1 und 126
- Klasse B zwischen 128.1 und 191.254
- Klasse C zwischen 192.1.1 und 223.254.254

- Um ein Netz zu adressieren, wird der Rechnerteil einer Adresse auf Null gesetzt, z.B. B-Netz der ETH: 129.132.0.0
- Um alle Rechner in einem Netz zu erreichen (Broadcast), wird der Rechnerteil auf 1 gesetzt, z.B. 129.132.255.255

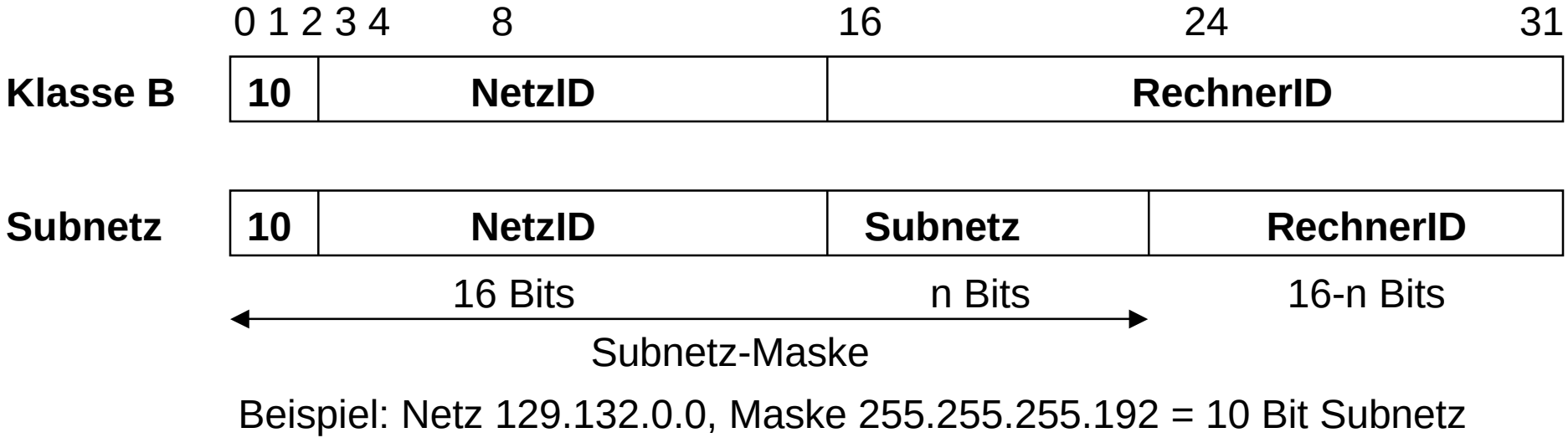
Adressierung im Internet: Spezielle Adressen

- Eine Null in einem Teil der Adresse bezeichnet per Konvention den lokalen Rechner bzw. das lokale Netz.
- Das Netz 127.0.0.0 ist der "lokale loopback Netz" eines Rechners. 127.0.0.1 ist die local loopback Adresse

Alles Null	**Lokaler Rechner**
Alles Null | **Rechner**	**Rechner auf lokalem Netz**
Alles Eins	**Beschränkter Broadcast (auf lokalem Netz)**
Netz | **Alles Eins**	**Gerichteter Broadcast für "Netz"**
127 | **Beliebig**	**Lokaler Loopback**

Adressierung im Internet: Namensautorität und Subnetze

- Die Adressautorität im Internet wird durch die zentrale Vergabe von Netzadressen durch das NIC in den USA ausgeübt.
- Um die Freiheit der lokalen Konfiguration zu erhöhen, und die Anzahl vergebener Netzadressen zu minimieren, ist die Verwendung lokaler Subnetz-Masken zur internen Unterteilung des Rechner-Teils der Adresse möglich.



Beispiel: Netz 129.132.0.0, Maske 255.255.255.192 = 10 Bit Subnetz

Adressierung im Internet: Merkpunkte

- Wenn ein Rechner an ein anderes Netz angehängt wird, muss seine Adresse geändert werden.
- Die Reihenfolge der Adressbytes ist im Standard festgelegt.
- Wenn z.B. ein C-Netz auf mehr als 255 Rechner wächst, müssen alle Rechner auf ein B-Netz migriert werden.
- Ein Rechner mit mehreren Anschlüssen an das Internet braucht mehrere Adressen, die auch verschiedene Routen implizieren.

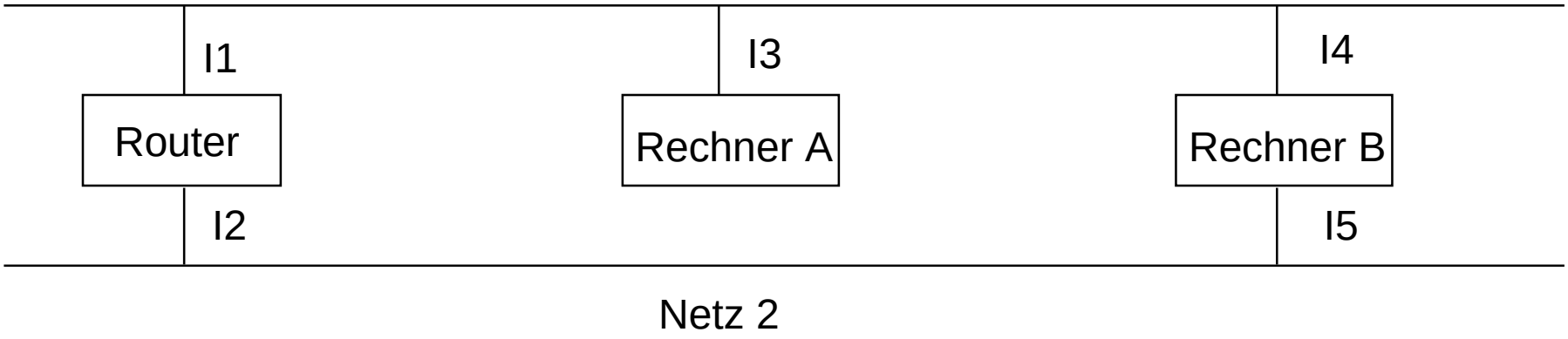


Abbildung von IP-Adressen auf physikalische Adressen

- Gegeben 2 Rechner, die am selben physikalischen Subnetz angeschlossen sind.
- Beide Rechner haben je eine IP-Adresse, und je eine physikalische Adresse bezüglich ihres gemeinsamen Netzes (z.B. eine Ethernet-Adresse).
- Will Rechner A Daten an Rechner B senden, so muss er anhand der IP-Adresse von Rechner B die Ethernet-Adresse von Rechner B herausfinden, um die Daten über das gemeinsame Netz zu senden.

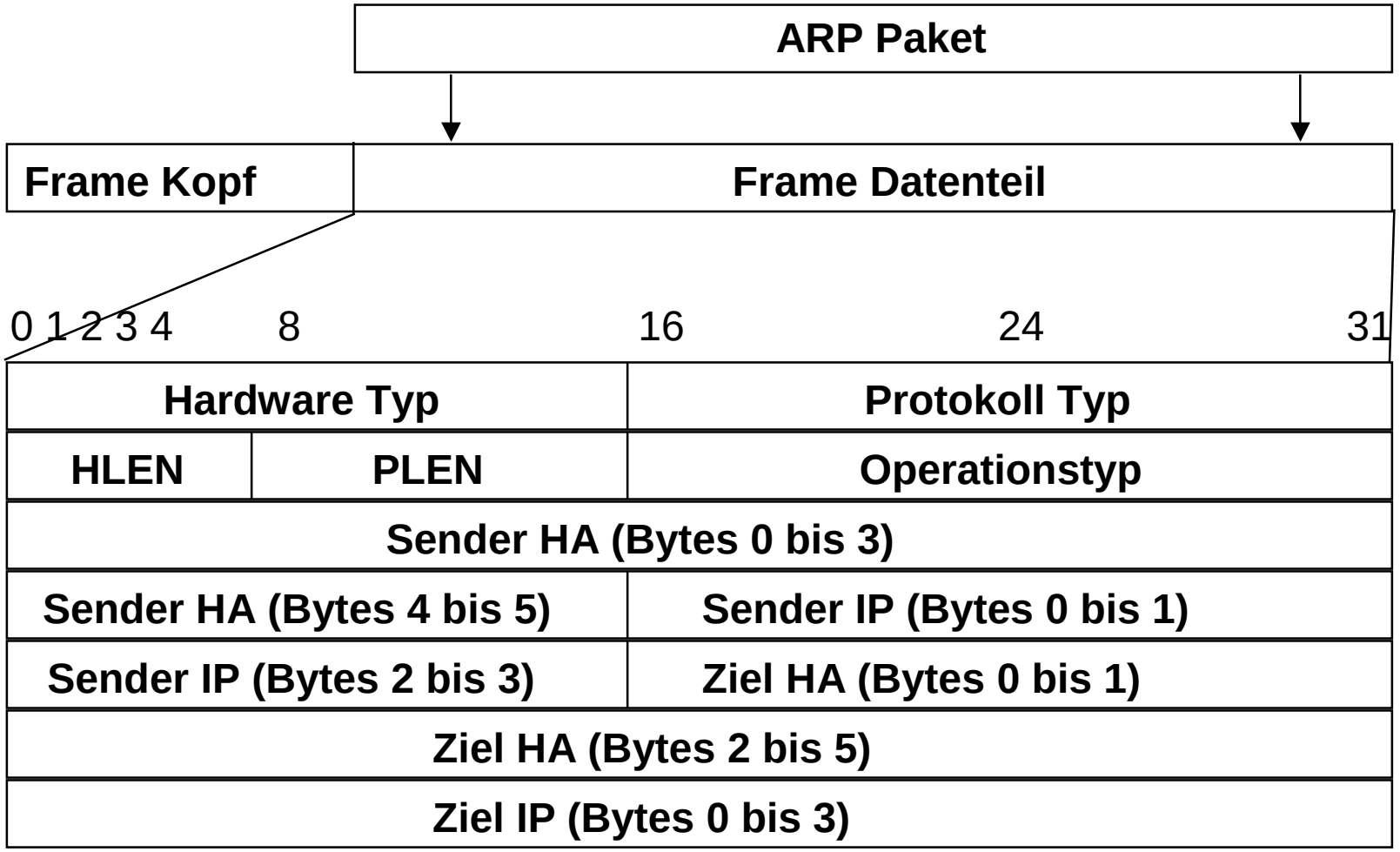
Wege zur Abbildung von IP-Adressen auf physikalische Adressen

- Direkte Abbildung durch Berechnung aus der IP-Adresse: dies ist nur in bestimmten Netzen möglich, und auch nur, solange das Adressierschema in beiden Adressräumen eingehalten wird.
- Suche der physikalischen Adresse in einem Verzeichnisdienst anhand der IP-Adresse.
- Dynamische Bindung durch Nachfragen auf dem lokalen Netz mittels des "Address Resolution Protocol": Rechner A sendet ein spezielles Broadcast-Paket auf das lokale Netz, in dem die IP-Adresse von Rechner B angegeben ist, und in dem nach der physikalischen Adresse von Rechner B gefragt wird. Rechner B füllt die gesuchte Adresse ein und sendet das Paket zurück.

Abbildung von IP-Adressen auf physikalische Adressen via ARP

- Gefundene Adressabbildungen werden von jedem Rechner für eine begrenzte Zeit in einem ARP-Cache lokal gehalten.
- Zur Optimierung wird die physikalische Adresse des fragenden Rechners dem Paket mitgegeben (Annahme von bidirektionalem Verkehr) und alle Rechner werten auch das Antwortpaket aus.

Aufbau eines ARP-Paketes



Beispiel der Aufdatierung des ARP-Cache

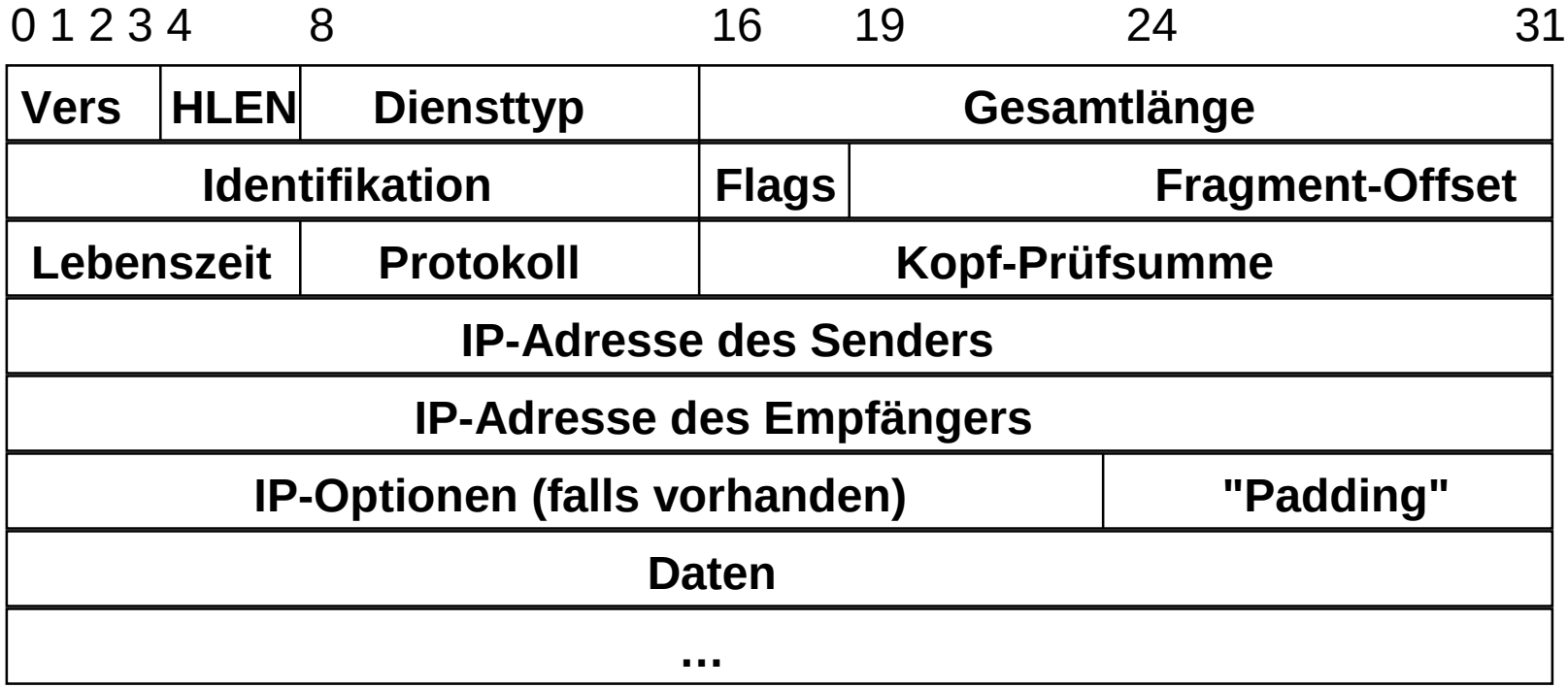
ktik0{lubich}[lubich]104> arp -a
ezci7-ifw (129.132.101.1) at aa:0:4:0:f3:e1 (6 Bytes, getrennt durch ":")
ktik5 (129.132.66.6) at 8:0:20:9:4a:e2
komsys.inf.ethz.ch (129.132.66.25) at 8:0:2b:2:ff:d5
komsys-gator (129.132.66.61) at 0:0:89:1:a0:f1
ktik0{lubich}[lubich]105> ping **ktik4**
ktik4 is alive
ktik0{lubich}[lubich]106> arp -a
ezci7-ifw (129.132.101.1) at aa:0:4:0:f3:e1
ktik4 (129.132.66.5) at 8:0:20:8:13:ea
ktik5 (129.132.66.6) at 8:0:20:9:4a:e2
komsys.inf.ethz.ch (129.132.66.25) at 8:0:2b:2:ff:d5
komsys-gator (129.132.66.61) at 0:0:89:1:a0:f1

Das Internet Protokoll (IP)

- Das IP-Protokoll implementiert einen verbindungslosen Daten-übermittlungsdienst (Datagram).

- definiert die "Basis-Dateneinheit" für das gesamte Internet
- implementiert die Weiterleitung von Paketen (packet forwarding)
- definiert einen Satz von Regeln für die Verwendung des Internet

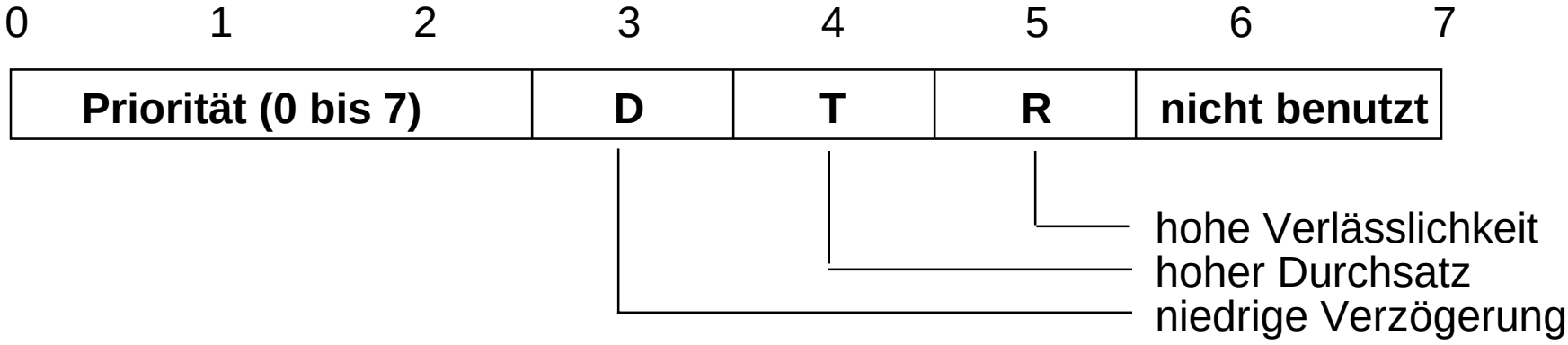
Aufbau eines IP-Paketes



Länge und Diensttypangabe in einem IP-Paket

- Das Feld HLEN gibt die Länge des IP-Kopfes an. Das einzige variabel lange Feld sind die IP-Optionen, die mittels des "Padding-Feldes" wieder auf eine fixe Länge gebracht werden. Eine typische Grösse (ohne Optionen) ist 20 Byte (HLEN = 5).

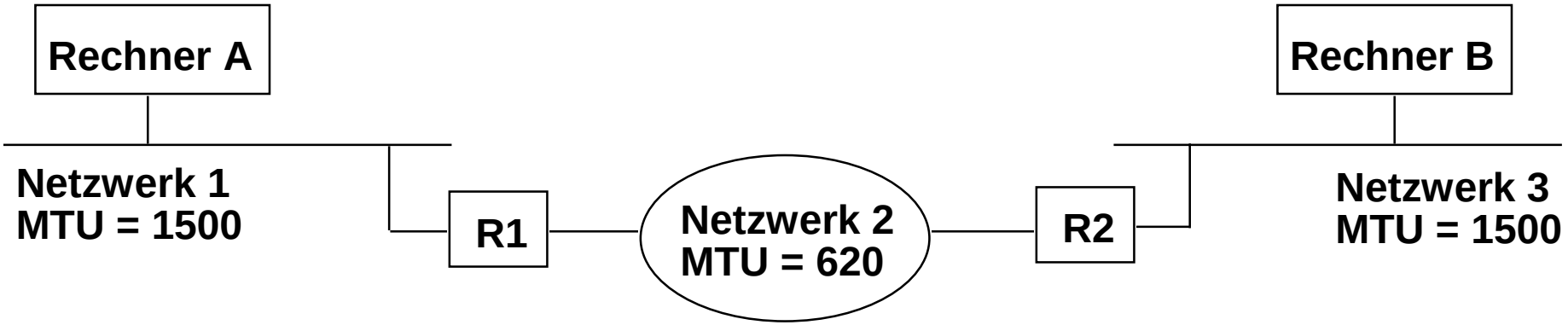
- Diensttyp:



Weitere Felder in einem IP-Paket: Fragmentierung

- Die Felder Identifikation, Flags und Fragment-Offset kontrollieren die Zerlegung zu langer Datenpakete in mehrere kleinere Pakete.
- Ethernet-Frames können maximal 1500 Bytes lang sein. Diese Maximalwerte werden als MTU (maximum transfer unit) bezeichnet.
- Ist ein IP-Paket grösser als die vorhandene MTU, muss das IP-Paket fragmentiert (zerlegt) werden.

Fragmentierung von IP-Paketen



Datagram-Kopf	1400 Bytes Daten		
Datagram-Kopf	600 Bytes Daten	600 Bytes Daten	200 Bytes Daten
Fragment-1-Kopf	600 Bytes Daten		
Fragment-2-Kopf	600 Bytes Daten		
Fragment-3-Kopf	200 Bytes Daten		

Flag "weitere Fragmente" gesetzt

Flag "weitere Fragmente" gesetzt

Weitere Felder in einem IP-Paket: Identifikation, Offset, Flags

- Das Feld "Identifikation" enthält eine eindeutige Nummer des ursprünglichen IP-Pakets.
- Der Fragment-Offset spezifiziert die Stelle im ursprünglichen IP-Paket, an dem das aktuelle Fragment eingesetzt werden muss.
- Mittels des Felds "Flags" kann Fragmentierung verboten werden, es wird auch zum Signalisieren weiterer Fragmente benutzt.
- Das Feld "Gesamtlänge" in einem Fragment bezieht sich auf die Länge des Fragmentes, nicht auf die Länge des IP-Pakets.

- Einmal fragmentierte Pakete werden erst beim Empfänger wieder zusammengesetzt (Nachteile: Zusatzlast und Gefahr von Verlust).

Weitere Felder in einem IP-Paket: Lebenszeit

- Das Feld "Lebenszeit" gibt an, wie lange (in Sekunden) ein Paket im Internet unterwegs sein darf, bevor es gelöscht wird. Beim Erstellen des Paketes wird eine Maximalzeit angegeben, die bei jeder Weiterleitung des Pakets dekrementiert wird.

- Wird ein Paket in einem Router verzögert, wird ein entsprechend höherer Wert abgezogen.

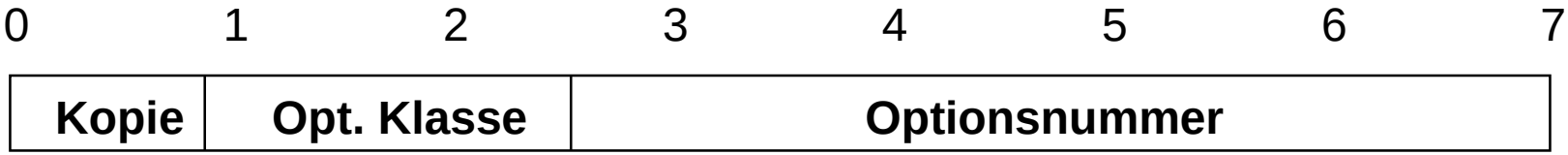
- Wird ein Paket wegen "Lebenszeit = 0" vor seiner Ankunft beim Empfänger gelöscht, muss das löschende System eine Fehlermeldung an den Urheber des Pakets zurücksenden.

Weitere Felder in einem IP-Paket: Protokoll, Prüfsumme, Adressen

- Das Feld "Protokoll" gibt an, welches hierarchisch über IP liegende Protokoll das Paket erzeugt hat, d.h. in welchem Format sich die Daten befinden.
- Das Feld Kopf-Prüfsumme dient der Datensicherung, bei der Bildung der Prüfsumme wird dieses Feld als "0" angenommen.
- Die Felder mit den IP-Adressen von Sender und Emfänger haben End-zu-End-Signifikanz, d.h sie werden nicht verändert, während das Paket durch das Internet transportiert wird.

Weitere Felder in einem IP-Paket: IP-Optionen

- IP-Optionen



- Das Flag "Kopie" gibt an, ob bei Fragmentierung die Optionen nur im ersten Fragment, oder in allen Fragmenten gesetzt werden.

- Optionsklasse

0	Normales Datengramm oder Netzwerk-Kontrolle
1	reserviert für zukünftige Benutzung
2	Fehlersuche und Messungen
3	reserviert für zukünftige Benutzung

Weitere Felder in einem IP-Paket: Optionale Elemente

- Optionen für Leitweglenkung und Zeitstempel
- Die Option "Wegaufzeichnung" ermöglicht die Protokollierung des Weges des Pakets durch das Netz im Optionsfeld des IP-Pakets.
- Die Option für die Wahl des Leitweges ermöglicht es dem Sender eines IP-Pakets, den Weg zum Empfänger zu diktieren. Diese Option wird meist zu Testzwecken benötigt.
- Die Zeitstempel-Option arbeitet ähnlich zur Option "Wegaufzeichnung", es wird jedoch ein zusätzlicher Zeitstempel angegeben. Weitere Flags steuern die Details der Angabe von Zeitstempeln.

Weiterführende Literatur

- Postel, Jon, "Internet Protocol - DARPA Internet Program Protocol Specification," RFC 791, Network Information Center, SRI International, Menlo Park, Calif., September 1981

- Leffler, McKusick, Karels, Quarterman, "The Design and Implementation of the 4.3BSD UNIX Operating System", Addison Wesley, Reading, MA, 1989, ISBN 0-201-06196-1

- W. Richard Stevens, "UNIX Network Programming", Prentice Hall, Engelwood Cliffs, NJ, 1990, ISBN 0-13-949876-1

- "man 4 ip"