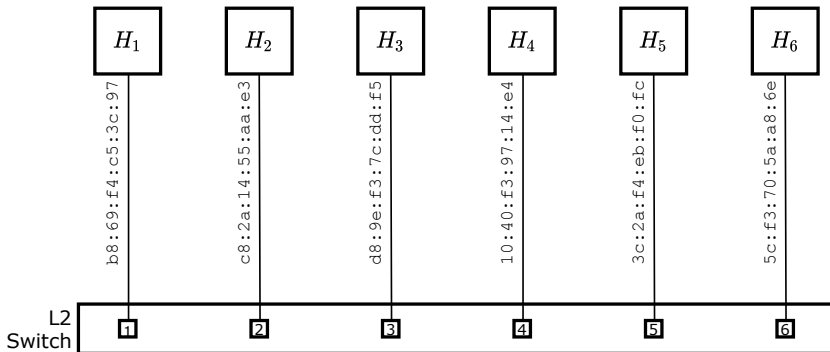


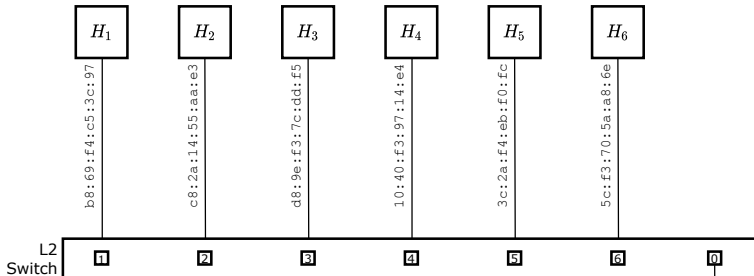
- picture example of L2-network not scaling
- picture intermediate device L2 interprets L2-addresses but does not change it
- L2 addresses are device-addresses (just a bunch of devices), L3 are "logical" (synthetic) addresses giving a meaning of localisation/group
- L3 allows every device to determine if a destination is *local* or *remote* (and only reachable via gateway)
- picture of L3 device interpreting L3 but not changing it
- picture of a whole lot of intermediate L3/router with separate L2 in-between

L2 revisited (1/2)



CAM Table	Port	MAC
1	1	b8:69:f4:c5:3c:97
2	2	c8:2a:14:55:aa:e3
3	3	d8:9e:f3:7c:dd:f5
4	4	10:40:f3:97:14:e4
5	5	3c:2a:f4:eb:f0:fc
6	6	5c:f3:70:5a:a8:6e

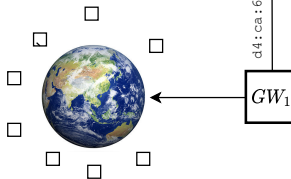
L2 revisited (2/2)



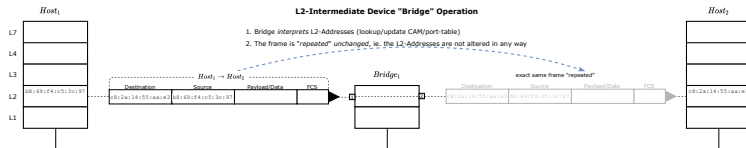
CAM Table

Port	MAC
1	<code>b8:69:f4:c5:3c:97</code>
2	<code>c8:2a:14:55:aa:e3</code>
3	<code>d8:9e:f3:7c:dd:f5</code>
4	<code>10:40:f3:97:14:e4</code>
5	<code>3c:2a:f4:eb:f0:fc</code>
6	<code>5c:f3:70:5a:a8:6e</code>
0	<code>d4:ca:6d:f8:6e:7e</code>
0	<code>d4:ca:6d:f8:6e:7e</code>
0	<code>00:0c:42:e9:25:57</code>
0	<code>d4:ca:6d:f8:6e:7e</code>
0	<code>c8:2a:14:55:aa:e3</code>
0	<code>00:08:9b:c1:17:fd</code>
0	<code>00:22:15:dd:59:16</code>
0	<code>00:08:9b:8c:c0:72</code>
0	<code>08:00:27:5e:78:6c</code>

... + all MAC-addresses in the world...



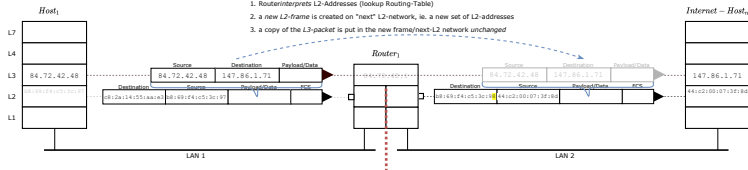
L2 Intermediate Device Operation (Bridge)



L3 Intermediate Device Operation (Router)

L3-Intermediate Device "Router" Operation

1. Router interprets L2-Addresses (lookup Routing-Table)
2. a new L2-frame is created on "next" L2-network, i.e. a new set of L2-addresses
3. a copy of the L3-packet is put in the new frame/next-L2 network unchanged

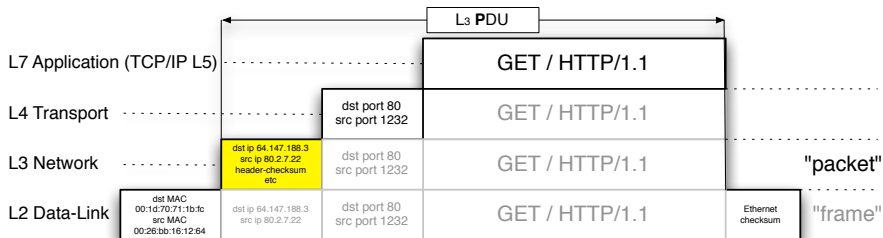


L3 Objectives

- Sie kennen die Unterschiede Layer-2 zu Layer-3 Adressierung
- Sie kennen die Strukturierung der Layer-3 Adressen in (Sub-) Netze
- Sie wissen, dass IP/Layer-3 ein *paketvermittelndes* Netz darstellt
- Sie verstehen das Konzept der *Netzmasken* und können es anwenden
- Sie wissen wie Layer-3 Pakete *geroutet*¹werden
- Sie wissen, dass IP/Layer-3 nach dem *best-effort* Prinzip arbeitet und die Zustellung der Pakete nicht garantiert ist

¹Wegleitung im Internet

L3 Stack



Header:

- *global* gültige *End-zu-End* (Geräte) Adressierung
- *Header-Prüfsumme*²
- Flags, Fragment, "Lebensdauer", Unterstützung für (einfache) Qualitätsicherung etc
- ... und natürlich das "upper-level-protocol"

²d.h. keine Prüfsumme über die SDU

L3 Factlets

- Layer-3 verbindet Systeme *End-to-End*³ – d.h. weltweit
- Layer-3 Adressen sind *strukturiert* in (Sub-) Netze⁴
- IP/Layer-3 ist ein *paketvermittelndes* Netz: die Nutzdaten werden paketisiert und gesondert übertragen
- verschiedene Layer-3 Netzwerke werden durch *Router*⁵ miteinander verbunden
- die Dateneinheit auf IP/Layer-3 ist das Paket “packet”⁶
- die Zustellung der Pakete auf IP/Layer-3 ist nicht garantiert⁷

³im Gegensatz zu Layer-2, das nur Teilstrecken/benachbarte Systeme verbindet

⁴wie z.B. das Telefon-Netzwerk

⁵oder “Gateway” (ungenauer)

⁶bei Layer-2 war das “frame”

⁷im Gegensatz zu Ethernet/L2 kann aber L3 bereits Fehlermeldungen zum Kommunikationspartner auslösen – *wenn* die Header-Checksum stimmt. . .

Unterschiede Layer-2 zu Layer-3 Adressierung

- Layer-2 ist eine *Geräte-Identifikation*⁸ ohne Struktur/Lokationsinformation
- ein Internet⁹ wäre mit Layer-2 Adressen nicht möglich, da die Bridge-Tabellen zu gross würden
- Layer-3/IP fasst mehrere Geräte in einem IP-(Sub-) Netz zusammen¹⁰ und erlaubt so eine effiziente lokalisierung
- über Layer-3/IP ist so eine weltweite lokalisierung/addressierung von *Endgeräten* möglich

⁸wie z.B. die AHV-Nummer

⁹weltweiter Netzwerkverbund

¹⁰wie bei den Telefonnummern die Landesvorwahl/Ortsvorwahl

Paketvermittelndes Netz

- Layer-3 unterteilt die Nutzdaten¹¹ in kleinere Einheiten und versendet diese gesondert über das Netz¹²
- Die Pakete können verloren gehen oder in anderer Reihenfolge¹³ am Ziel ankommen – die Wiederherstellung der Nutzdaten ist Aufgabe der oberen Layer
- ein paketvermittelndes Netz ist Fehlerresistent¹⁴
- ... und erlaubt eine effiziente Auslastung¹⁵ der Ressourcen

¹¹z.B. Webseiten oder Bild- und Tondaten

¹²wie auch schon auf Layer-2

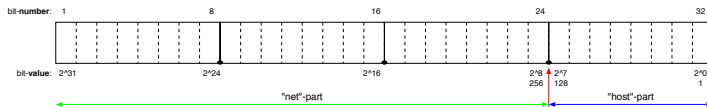
¹³oder auch *verdoppelt* werden

¹⁴solange die oberen Layer für die garantierte Zustellung aufkommen. Der Ausfall einzelner Verbindungen/L2 im Internet ist üblicherweise kein Problem

¹⁵im Gegensatz zu leitungsvermittelnden Netzen – wo die Reservierung der Bandbreite unabhängig der tatsächlichen Ausnutzung erfolgt

Struktur der IP-Adressen “(Sub-) Netzwerke”¹⁹

- IP-Adressen bestehen aus 32-Bit und werden üblicherweise im “dotted-decimal”¹⁶ Format notiert: 194.41.161.1



- die 32-Bit werden in *Netz*- und *Host*-Anteil unterteilt¹⁷
 - ▶ Netz: entspricht etwa Landes-, Ortsvorwahl im Telefonnetz
 - ▶ Host: entspricht etwa Nummer des Telefonapparates ohne Vorwahl
- Das “routing”¹⁸ wird aufgrund des Netz-Anteils bestimmt
 - ▶ ... wie beim Telenfonnetz...

¹⁶ 4-Bytes im Dezimalformat 0-255 getrennt durch “.”

¹⁷ manchmal zusätzlich in Sub-Netz-Anteil

¹⁸ Wegleitung

¹⁹ im Folgenden werden “Subnetze” und “Netze” weitgehend Synonym verwendet

Netz-Anteil

- der Netz-Anteil kann mittels *Netzmasken* oder *Präfixlängen* angegeben werden:
 - ▶ bit-and Maske im “dotted-decimal” Format: **255.255.255.0**, binär: **11111111.11111111.11111111.00000000**
 - ▶ Präfixlänge (Anzahl gesetzter=“1” Bits²⁰ in der Maske): **/24**
- Netzzugehörigkeit: wird mittels der logischen “Bitwise-And” Operation durchgeführt
 - ▶ IP-Adresse in binäres Format umrechnen, eg: $192.168.1.5_{10} \rightarrow 11000000.10101000.00000001.00000101_2$
 - ▶ IP-Maske in binäres Format verwandeln, eg: $255.255.255.0_{10} \rightarrow 11111111.11111111.11111111.00000000_2$
 - ▶ “bitwise-and” Verknüpfung von IP-Adresse und -Maske, eg:
 $11000000.10101000.00000001.00000101_2 \& 11111111.11111111.11111111.00000000_2 =$
 $11000000.10101000.00000001.00000000_2 \rightarrow 192.168.1.0_{10}$
 - ▶ ... die so gewonnene Adresse wird als *Netzbasisadresse* bezeichnet und kann *nicht* für ein Endgerät/Node verwendet werden
- der Host-Anteil kann auf gleiche Weise gewonnen werden, wenn die Maske zuerst *invertiert*²¹ wird

²⁰ “consecutive 1-bits from MSB”

²¹ “one’s-complement”

Interlude

- finden Sie die IP-Adresse und die Netzmaske Ihres Laptops²²
- bestimmen Sie aus diesen Informationen die *Netzbasisadresse* und den Host-Anteil
- bestimmen Sie ob die beiden IP-Adressen 192.168.2.126 und 192.168.2.130 bei Verwendung einer Netzmaske (für beide) von 255.255.255.128 im selben Netz liegen
- bestimmen Sie die Anzahl möglicher²³ IP-Adressen bei Verwendung der Netzmasken 255.255.255.0 und 255.255.254.0
- berechnen Sie die “längste”²⁴ Netzmaske wenn rund 2000 Adressen im selben Netz benötigt werden

²²ifconfig/ipconfig, netstat -r

²³Gesamtanzahl minus zwei: broadcast und base

²⁴d.h. das eben gerade ausreichende kleinste Netzwerk

Aufteilung des Adressbereichs

- bis ca 1993 “chaotisch”:

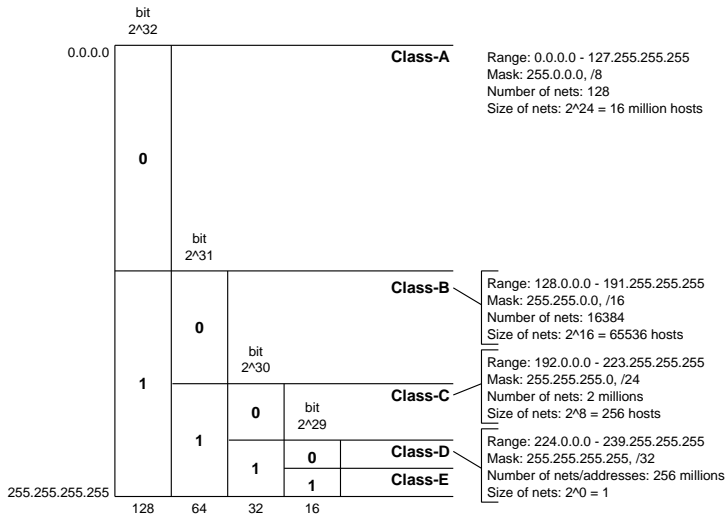
<http://www.iana.org/assignments/ipv4-address-space/ipv4-address-space.xml>

- seither (CIDR) in geographisch/institutionell hierarchischer Weise über “Unterverteiler” RIR²⁵, private Provider²⁶

²⁵ regional-internet-registry, z.B. <http://www.ripe.net/>

²⁶ z.B. 212.x.x.x und 213.x.x.x sind beide “Europa” – ähnlich wie “Landesvorwahl”

IP-Adressbereich als Klassen "the olde man" 1/2



IP-Adressbereich als Klassen “*he olde man*” 2/2

- es werden mit *ID-Bits* identifizierte Adressbereiche für grosse, mittlere und kleine Netze gebildet
- ... “it seemed a good idea in the 80’”
- Problem-1: nur 128 Netze belegen die Hälfte des Addressraums (aber können 16 Millionen Hosts aufnehmen)
- Problem-2: keine topographisch/hierarchische²⁷ Aufteilung
- Problem-3²⁸: Router müssen im schlechtesten Fall etwas über 2 *Millionen* Netze kennen

²⁷wie z.B. das Telefonnetz

²⁸big, fat...

CIDR “Classless Inter-Domain Routing” (the new way) 1/2

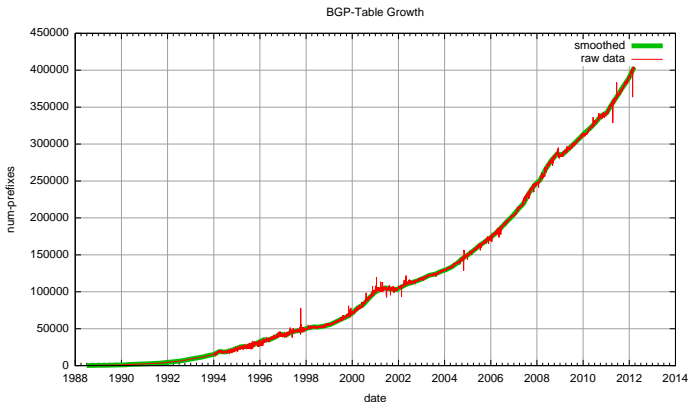
- Neustrukturierung²⁹ des Adressbereichs
- Klassen werden *nicht* mehr beachtet
- Bildung von *Supernetzen* für effizientes hierarchies routing³⁰
→ kleine Routing-Tables
- Bessere Ausnutzung des Adressbereichs durch Aufteilung der A- und B-Klassen in kleinere Einheiten
- Notation als *Präfixlänge*³¹: 255.255.240.0→/20

²⁹soweit möglich... Bereits zugewiesene Netze können nicht entfernt werden

³⁰eg. 212/7→Europa

³¹“slash”-Notation

CIDR “Classless Inter-Domain Routing” (the new way) 2/2



```
#!/usr/bin/env gnuplot -persist
set grid
set xdata time
set timefmt "%s"
set format x "%Y"
set title "BGP-Table Growth"
set ylabel "num-prefixes"
set xlabel "date"
plot "BGP-Table Growth" using lines title "smoothed" with lines,
```

... noch mehr Spass mit Bits

- Berechnen Sie zu der CIDR-Präfixlänge /20 die entsprechende Netzmaske
- ... und umgekehrt zu der Netzmaske 255.255.192.0 die Präfixlänge
- wieviele mögliche IP-Adressen können in /21 untergebracht werden?
- Bestimmen Sie ob die beiden IP-Adressen 172.17.71.5/23 und 172.17.70.240 im selben Netz liegen

Routing und Routing-Table

- die Wegleitung “routing” im Internet:
 - ▶ wird bei jedem Router³² neu entschieden und dann an den nächsten Router oder an den Zielhost zugestellt
 - ▶ entschieden wird der “next-hop” über die *Routing-Tabelle*³³
- die *Routing-Tabelle* enthält (mindestens):
 - ▶ *Ziel-Netz* (dest-net)
 - ▶ *Ziel-Maske* oder *Ziel-Präfixlänge* (dest-mask, prefixlength)
 - ▶ *Nächster-Router* (next-hop oder gateway)
- der *Routing-Vorgang* (forwarding): findet den “next-hop” für eine gegeben/empfangene IP-Ziel-Adresse (target-ip)
 - ▶ RT-Eintrag “passt”, wenn:
(target-ip & dest-mask_i) = dest-net_i, dann wird das Paket an next-hop_i weitergeleitet
 - ▶ die RT wird nach Präfixlänge in *absteigender* Reihenfolge konsultiert
 - ▶ die “Default-Route” ist 0.0.0.0/0

³² “hop-by-hop” und “next-hop”

³³ im Folgenden “routing-table” oder RT – eigentlich ist das die FIB

References

- http://en.wikipedia.org/wiki/IPv4_subnetting_reference
- http://en.wikipedia.org/wiki/Classless_Inter-Domain_Routing
- Route-servers: <http://www.traceroute.org/#Route%20Servers>
- BGP-routing-table growth: <http://bgp.potaroo.net/as2.0/bgp-active.html>, andere lustige Informationen: <http://bgp.potaroo.net/as2.0/>
- CIDR: <http://books.google.ch/books?id=axiW1d8GosIC&lpg=PA125&pg=PA101#v=onepage&q=&f=false>
- IP-address landscape: <http://xkcd.com/195/>