Empfohlene Links

- Google Video-Kurs (5h!) https://www.youtube.com/watch?v=QKfk7YFILws
- Netz-Mafia Grundlagen: http://www.netzmafia.de/skripten/netze/index.html
- Netz-Mafia Internet: http://www.netzmafia.de/skripten/internet/index.html

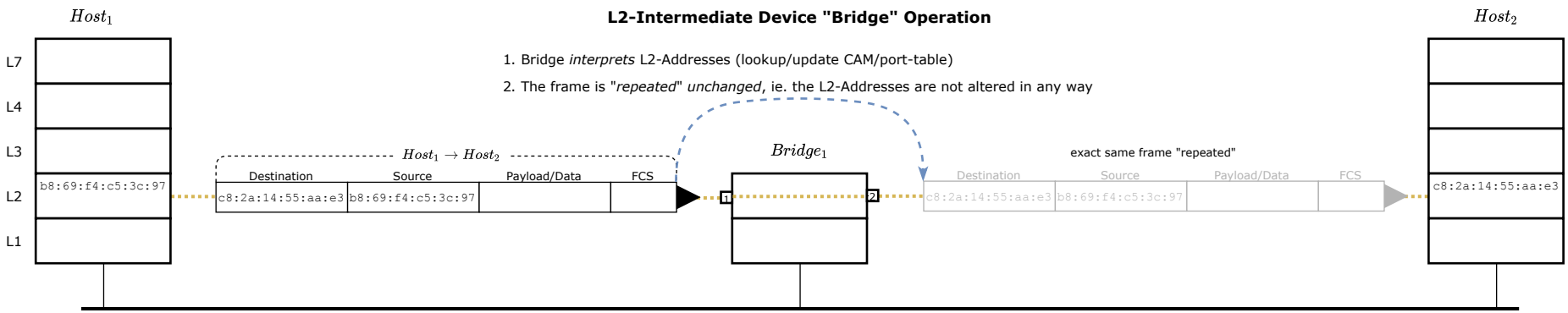
- routing is done on a next-hop basis
- ARP
- L3 allows every device to determine if a destination is *local* or *remote* (and only reachable via gateway)
- picture of a whole lot of intermediate L3/router with seperate L2 in-between

L3: Lenrziele

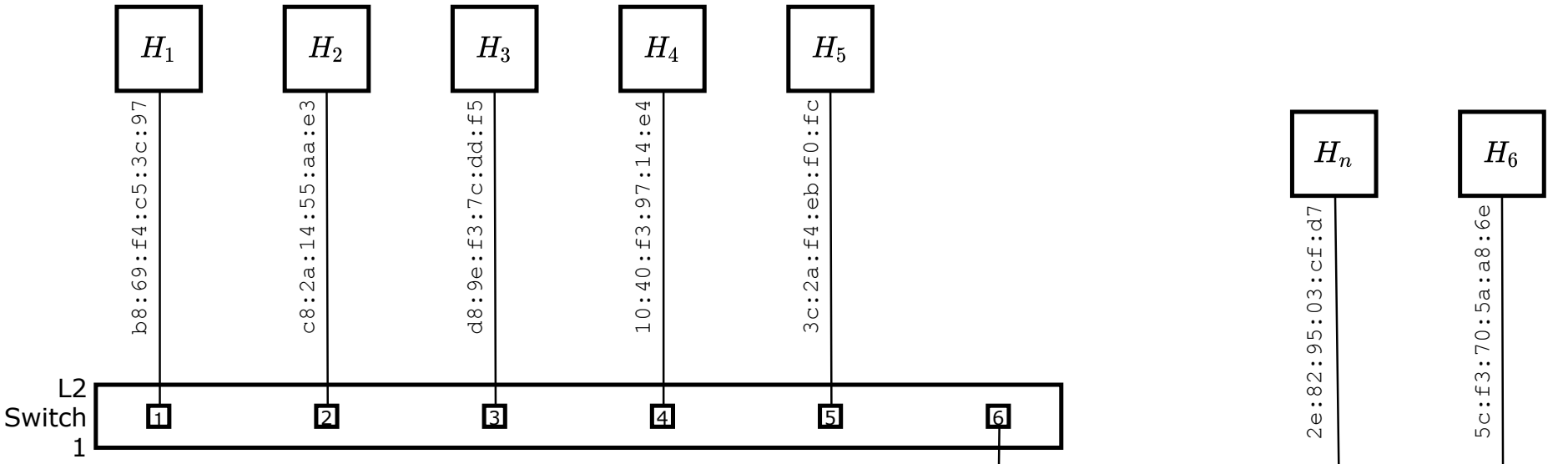
- Sie kennen die Unterschiede Layer-2 zu Layer-3 Adressierung
- Sie kennen die Strukturierung der Layer-3 Adressen in (Sub-) Netze
- Sie wissen, dass IP/Layer-3 ein *paketvermittelndes* Netz darstellt
- Sie verstehen das Konzept der *Netzmasken* und können es anwenden
- Sie wissen wie Layer-3 Pakete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

[1] Wegleitung im Internet

L2 revisited: Intermediate Device Operation (Bridge) (1/4)



L2 revisited: LAN-extension (same LAN) (2/4)



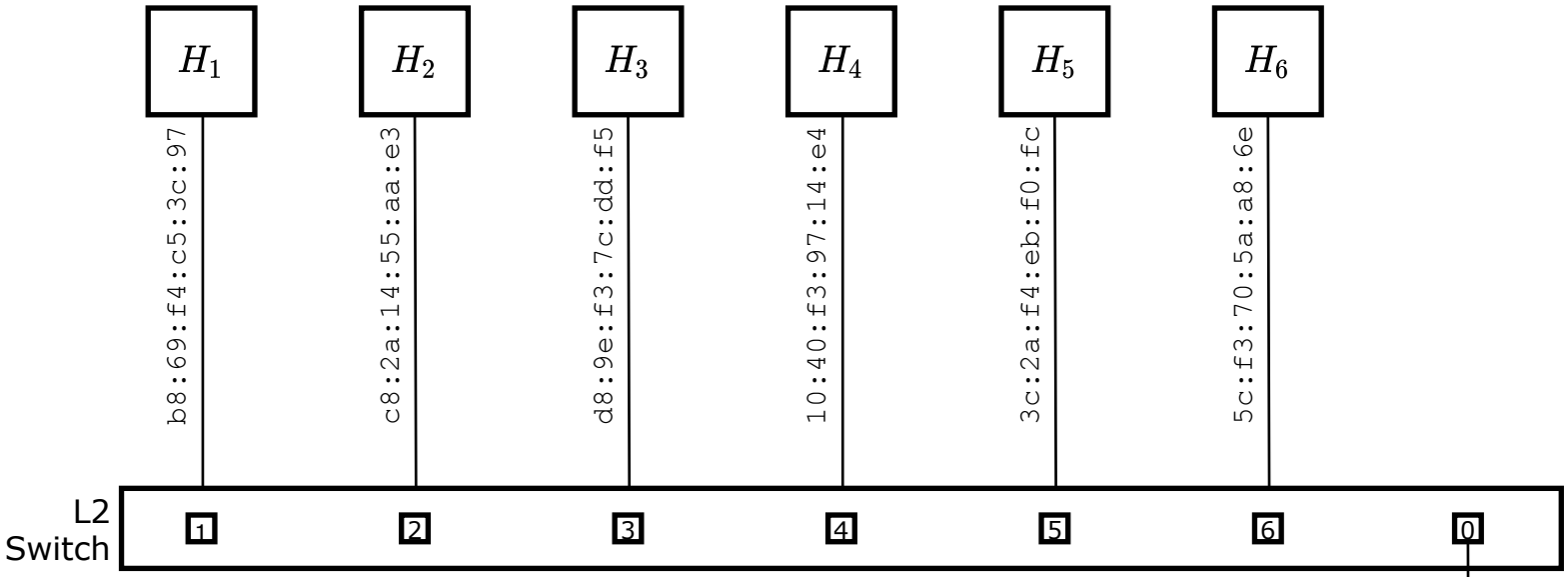
CAM Table 1

Port	MAC
1	b8:69:f4:c5:3c:97
2	c8:2a:14:55:aa:e3
3	d8:9e:f3:7c:dd:f5
4	10:40:f3:97:14:e4
5	3c:2a:f4:eb:f0:fc
6	5c:f3:70:5a:a8:6e
6	2e:82:95:03:cf:d7

CAM Table 2

Port	MAC
1	b8:69:f4:c5:3c:97
1	c8:2a:14:55:aa:e3
1	d8:9e:f3:7c:dd:f5
1	10:40:f3:97:14:e4
1	3c:2a:f4:eb:f0:fc
4	5c:f3:70:5a:a8:6e
3	2e:82:95:03:cf:d7

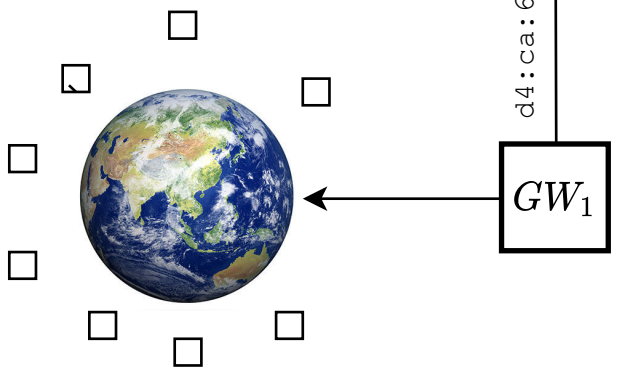
L2 revisited: "Internet" with L2? (3/4)



CAM Table

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

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



L2 revisited: L2-"Internet" Problems (4/4)

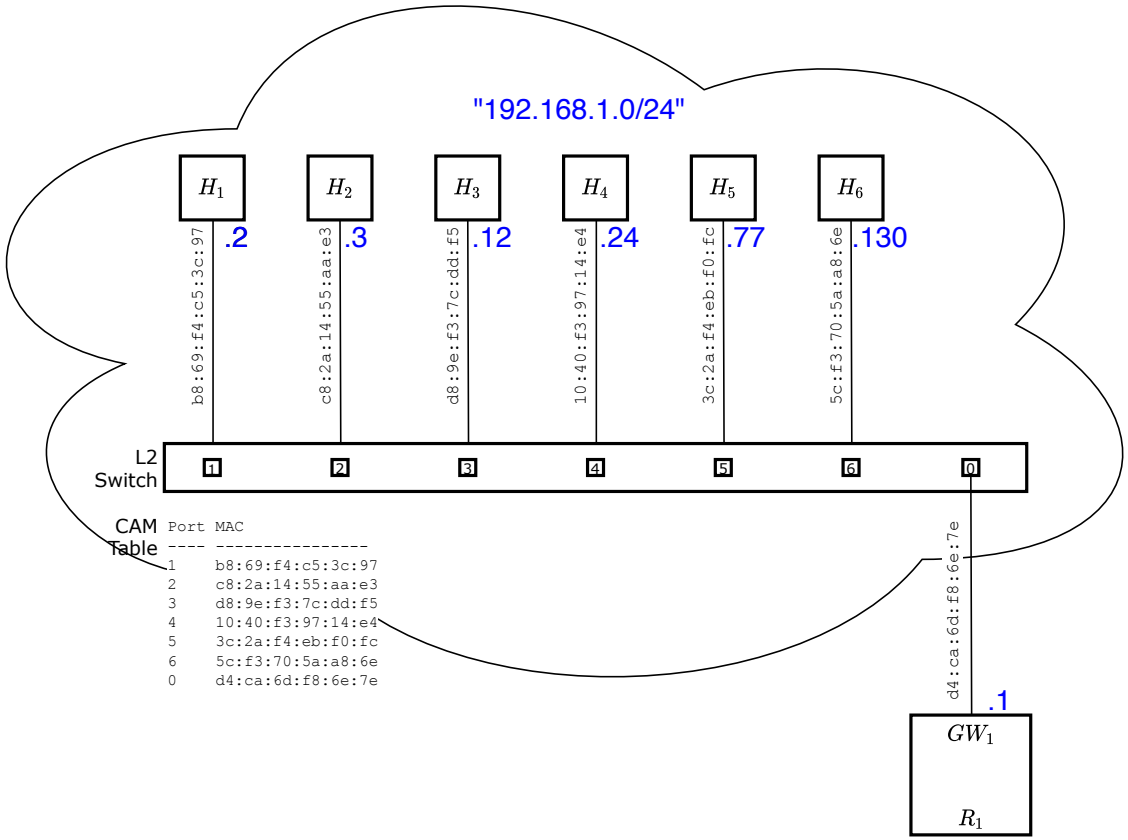
- die CAM/Port-Tabelle wächst bei Erweiterung des Netzwerks potentiell bis zur Anzahl Hosts
- ein weltweites Internet ist so nicht realisierbar: die CAM-Tabelle würde alle potentiellen Ziele im Internet unter dem "Uplink"-Port speichern
- einfach alles unbekannte per "Flooding" auf den Uplink schicken funktioniert ebenfalls nicht:
  - ein weltweites Internet würde nur duch das Flooding ausgelastet
  - zudem könnte alle Hosts mithören (der "privater Link" Vorteil geht verloren)

MAC-Adressen sind *Geräteadressen* (Device-Address) und haben keine strukturelle Lokalisierung/Gruppenzusammengehörigkeit

$$\text{LAN} = \text{"Just a bunch of devices"}$$

L2: the need for L3 :) (1/2)

- Hosts/Geräte müssen in einer lokalisierten Gruppe zusammengefasst werden: “logische” (abstrakte) Adressierung
- *forwarding*/Weiterleitung Aufgrund der Gruppe (“Netz”) und nicht der einzelnen Geräteadressen → viel kleinere “Routingtabellen”



L2: the need for L3 (2/2)

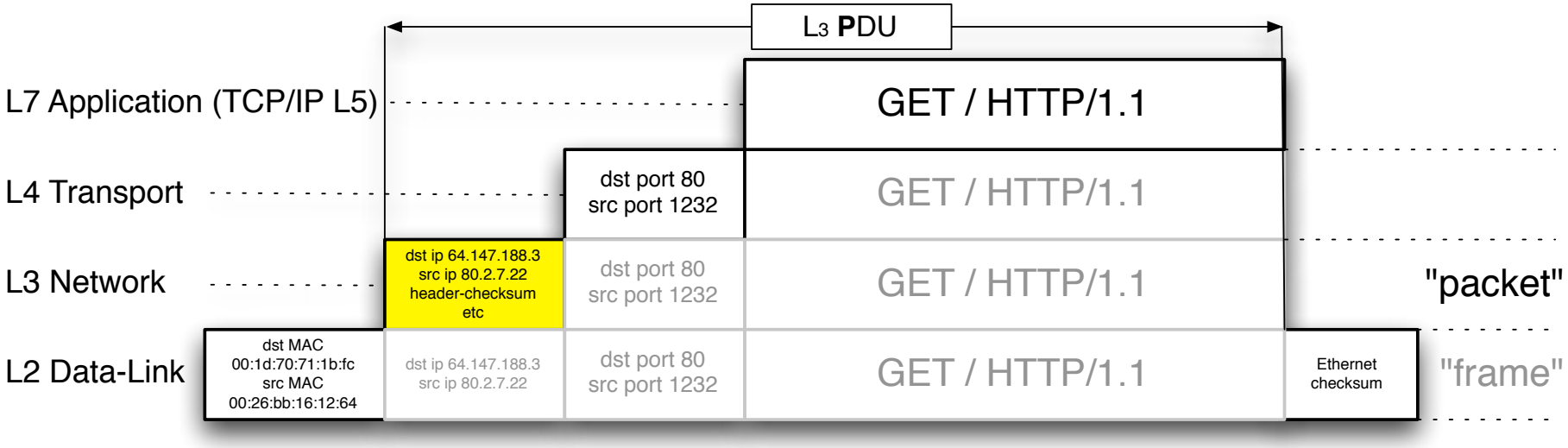
- Die IP-Adressierung ist vergleichbar mit den Telefonnummern: zuerst die globale Vorwahl, Ortsvorwahl und dann die Apparatennummer
  - Die Adressen sind so strukturiert, dass die "Geräte" /Hosts *rechts* nummeriert werden
  - die "Vorwahl" /Gruppe ist *links* gehalten

Die Adressen sind in der Regel *ortsgebunden* (localized) und werden auf die Geräte *konfiguriert*[a]

[a] im Gegensatz zu den MAC/L2-Adressen, die fest eingestellt sind. Das kann automatisch (DHCP) oder manuell erfolgen

- ein L2-Netzwerk kann beliebig viele L3-Gruppen (Netze) enthalten[2], das stiftet aber in der Regel Verwirrung :)

[2] d.h. L2 funktioniert unabhängig von L3



Header:

- *global* gültige *End-zu-End* (Geräte) Adressierung
- *Header*-Prüfsumme[3]
- Flags, Fragment, "Lebensdauer", Untersützung für (einfache) Qualitätsicherung etc
- . . . und natürlich das "upper-level-protocol"

L3 Factlets

- Layer-3 verbindet Systeme *End-to-End* [4] – d.h. weltweit
- Layer-3 Adressen sind *strukturiert* in (Sub-) Netze[5]
- IP/Layer-3 ist ein *paketvermittelndes* Netz: die Nutzdaten werden paketisiert und gesondert übertragen
- verschiedene Layer-3 Netzwerke werden durch *Router*[6] miteinander verbunden
- die Dateneinheit auf IP/Layer-3 ist das Paket "packet" [7]
- die Zustellung der Pakete auf IP/Layer-3 ist nicht garantiert[8]

[4] im Gegensatz zu Layer-2, das nur Teilstrecken/benachbarte Systeme verbindet

[5] wie z.B. das Telefon-Netzwerk

[6] oder "Gateway" (ungenauer)

[7] bei Layer-2 war das "frame"

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

Paketvermittelndes Netz

- Layer-3 unterteilt die Nutzdaten[9] in kleinere Einheiten und versendet diese gesondert über das Netz[10]
- Die Pakete können verloren gehen oder in anderer Reihenfolge[11] am Ziel ankommen – die Wiederherstellung der Nutzdaten ist Aufgabe der oberen Layer
- ein paketvermittelndes Netz ist Fehlerresistent[12]
- . . . und erlaubt eine effiziente Auslastung[13] der Resourcen

[9] z.B. Webseiten oder Bild- und Tondaten

[10] wie auch schon auf Layer-2

[11] oder auch *verdoppelt* werden

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

[13] im Gegensatz zu leitungsvermittelnden Netzen – wo die Reservierung der Bandbreite unabhängig der tatsächlichen Ausnutzung erfolgt

Unterschiede Layer-2 zu Layer-3 Adressierung

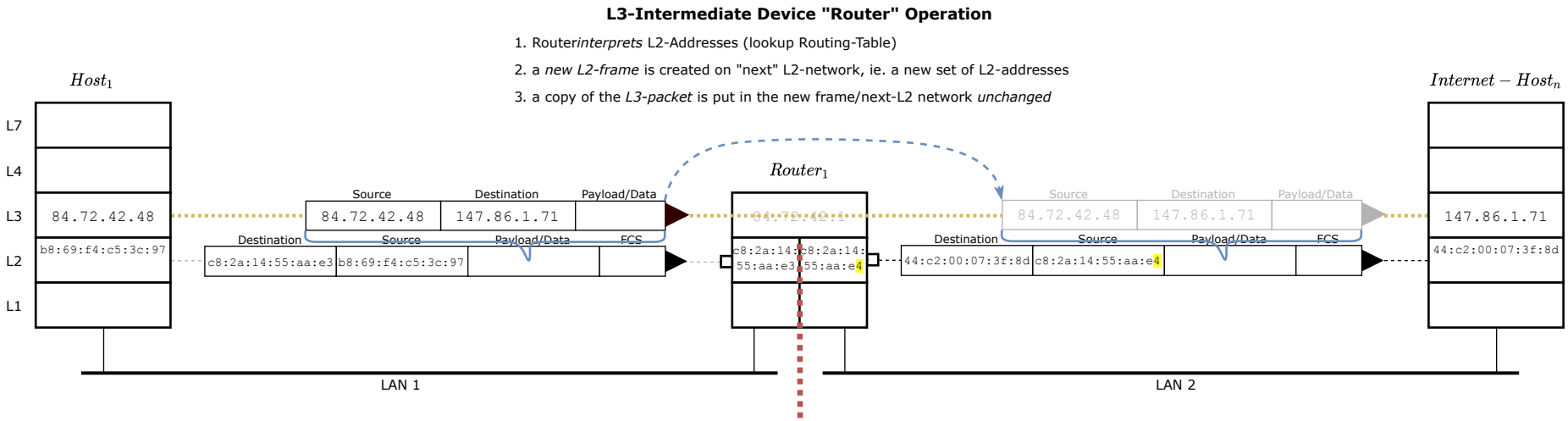
- Layer-2 ist eine *Geräte-Identifikation*[14] *ohne* Struktur/Lokationsinformation
- ein Internet[15] wäre mit Layer-2 Adressen nicht möglich, da die Bridge-Tabellen zu gross würden
- Layer-3/IP fasst mehrere Geräte in einem IP-(Sub-) Netz zusammen[16] und erlaubt so eine effiziente lokalisierung
- über Layer-3/IP ist so eine weltweite lokalisierung/addressierung von *Endgeräten* möglich

[14] wie z.B. die AHV-Nummer

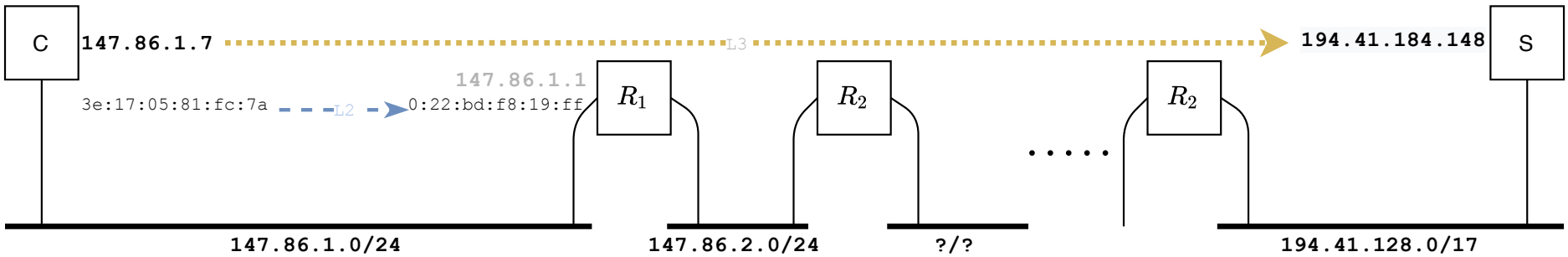
[15] weltweiter Netzwerkverbund

[16] wie bei den Telefonnummern die Landesvorwahl/Ortsvorwahl

L3 Intermediate Device Operation (Router)



L3 end-to-end


C
147.86.1.7
L3
194.41.184.148
S
147.86.1.1
3e:17:05:81:fc:7a
L2
0:22:bd:f8:19:ff
R_1
R_2
R_2
147.86.1.0/24
147.86.2.0/24
?/?
194.41.128.0/17

Struktur der IP-Adressen (1/3)

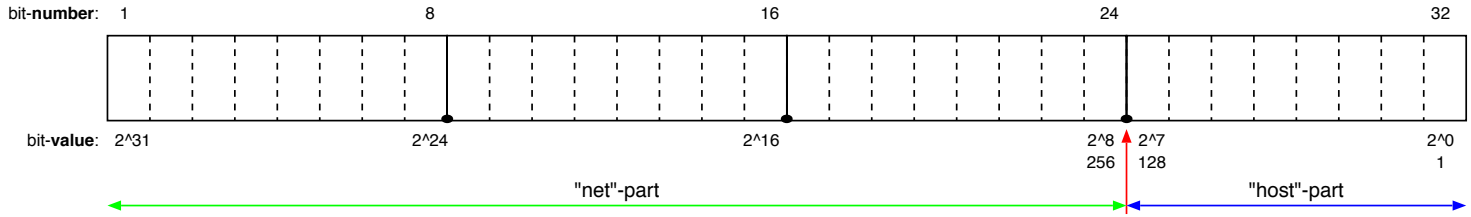
- IP-Adressen werden wie Telefonnummern von links nach rechts "spezifischer":
  - links ist die globale Einordnung: "Vorwahl" (prefix)
  - rechts ist der einzelne Anschluss/Apparat
- "Routing" basiert nur auf dem "Vorwahl"-Teil[17], es muss eine effiziente[18] Methode zur "Extraktion" des Prefix gefunden werden

[17] d.h. weit entfernte Router brauchen nur den Prefix und nicht die vollständige Adresse

[18] weil *jedes* Paket aufs Neue "geroutet" wird

Struktur der IP-Adressen: Prefix/Netz-Anteil (2/3)

- IPv4-Adressen sind 32-Bit breit und werden üblicherweise im "dotted-decimal"[19] Format notiert: 194.41.161.1



- die 32-Bit werden in *Netz*- und *Host*-Anteil unterteilt[20]
  - Netz: entspricht etwa Landes-, Ortsvorwahl im Telefonnetz
  - Host: entspricht etwa Nummer des Telefonapparates ohne Vorwahl
- Das "routing"[21] wird aufgrund des Netz-Anteils bestimmt
  - ...wie beim Telenfonnetz...
- die Aufteilung ist im Gegensatz zum Telefonsystem in weiten Grenzen anpassbar[22]

[19] 4-Bytes im Dezimalformat 0-255 getrennt durch "."

[20] manchmal zusätzlich in Sub-Netz-Anteil

[21] Wegleitung

[22] d.h. der Prefix hat nicht eine fest vorgegebene Grösse

Struktur der IP-Adressen: Prefix-Extraktion (3/3)

- der Netz-Anteil kann mittels *Netzmasken* oder *Präfixlängen* angegeben werden:
  - bit-and Maske im "dotted-decimal" Format: 2**55.255.255.0**, binär: 1**1111111.11111111.11111111.00000000**
  - Präfixlänge (Anzahl gesetzter="1" Bits[23] 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: $11000000.10101000.00000001.00000101_2$ **&** $11111111.11111111.11111111.00000000_2 =$ $11000000.10101000.00000001.00000\mathbf{000}_2 \rightarrow 192.168.1.\mathbf{0}_{10}$
  - . . . 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*[24] wird

[23] "consecutive 1-bits from MSB"

[24] "one's-complement"

Interlude

- finden Sie die IP-Adresse und die Netzmaske Ihres Laptops[25]
- 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 möglicher[26] IP-Adressen bei Verwendung der Netzmasken `255.255.255.0` und `255.255.254.0`
- berechnen Sie die "längste"[27] Netzmaske wenn rund 2000 Adressen im selben Netz benötigt werden

[25] `ifconfig`/`ipconfig`, `netstat -r`

[26] Gesamtanzahl minus zwei: broadcast und base

[27] d.h. das eben gerade ausrechende kleinste Netzwerk

Aufteilung des Adressbereichs

- bis ca 1993 “chaotisch”:

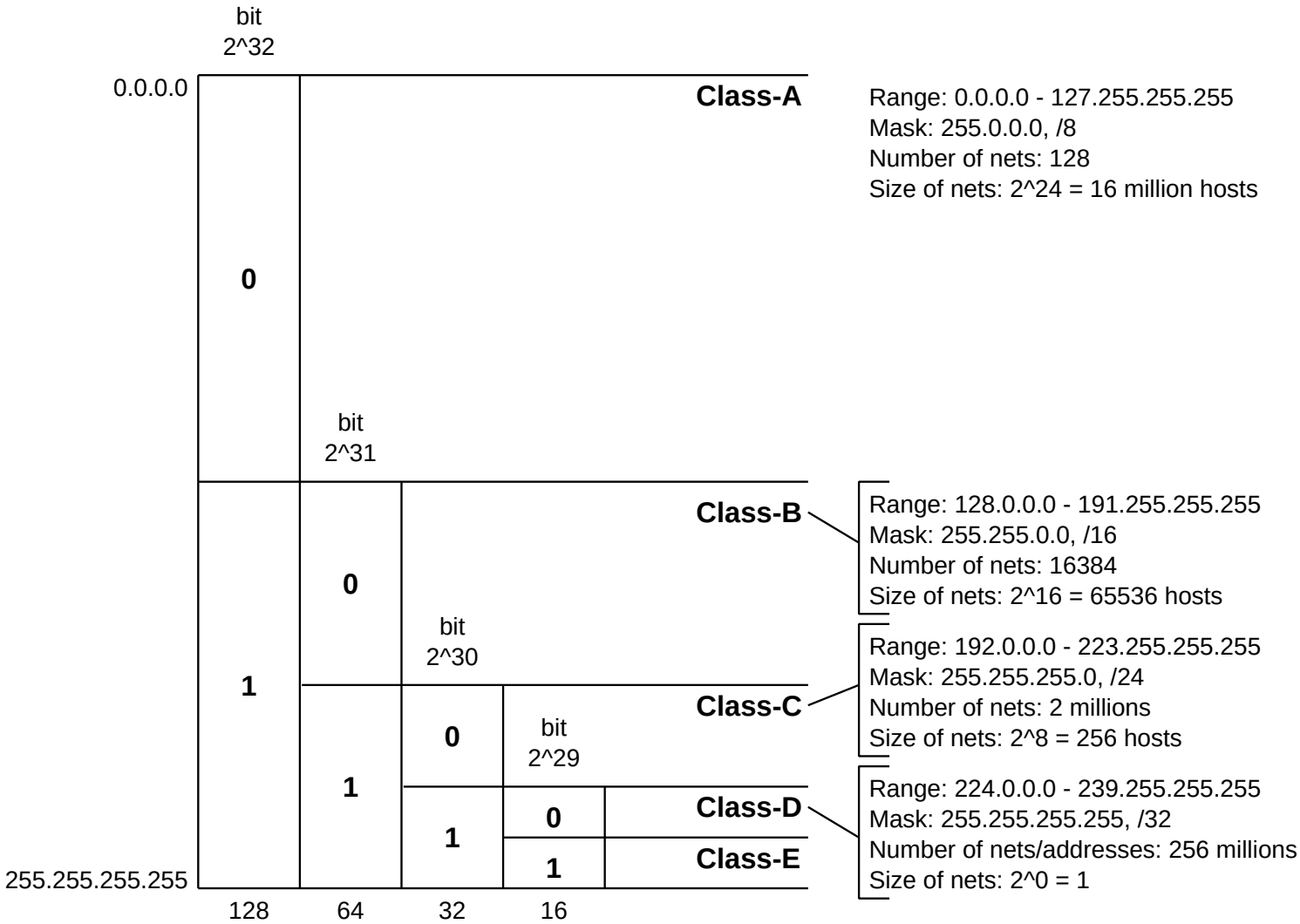
  http://www.iana.org/assignments/ipv4-address-space/ipv4-address-space.xml

- seither (CIDR) in geographisch/institutionell hierarchischer Weise über “Unterverteiler” RIR[28], private Provider[29]

[28] regional-internet-registry, z.B. http://www.ripe.net/

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

IP-Adressbereich als Klassen "ye olde way" 1/2



IP-Adressbereich als Klassen "ye olde way" 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/hierarchische[30] Aufteilung
- Problem-3[31]: Router müssen im schlechtesten Fall etwas über *2 Millionen* Netze kennen

[30] wie z.B. das Telefonnetz

[31] big, fat. . .

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

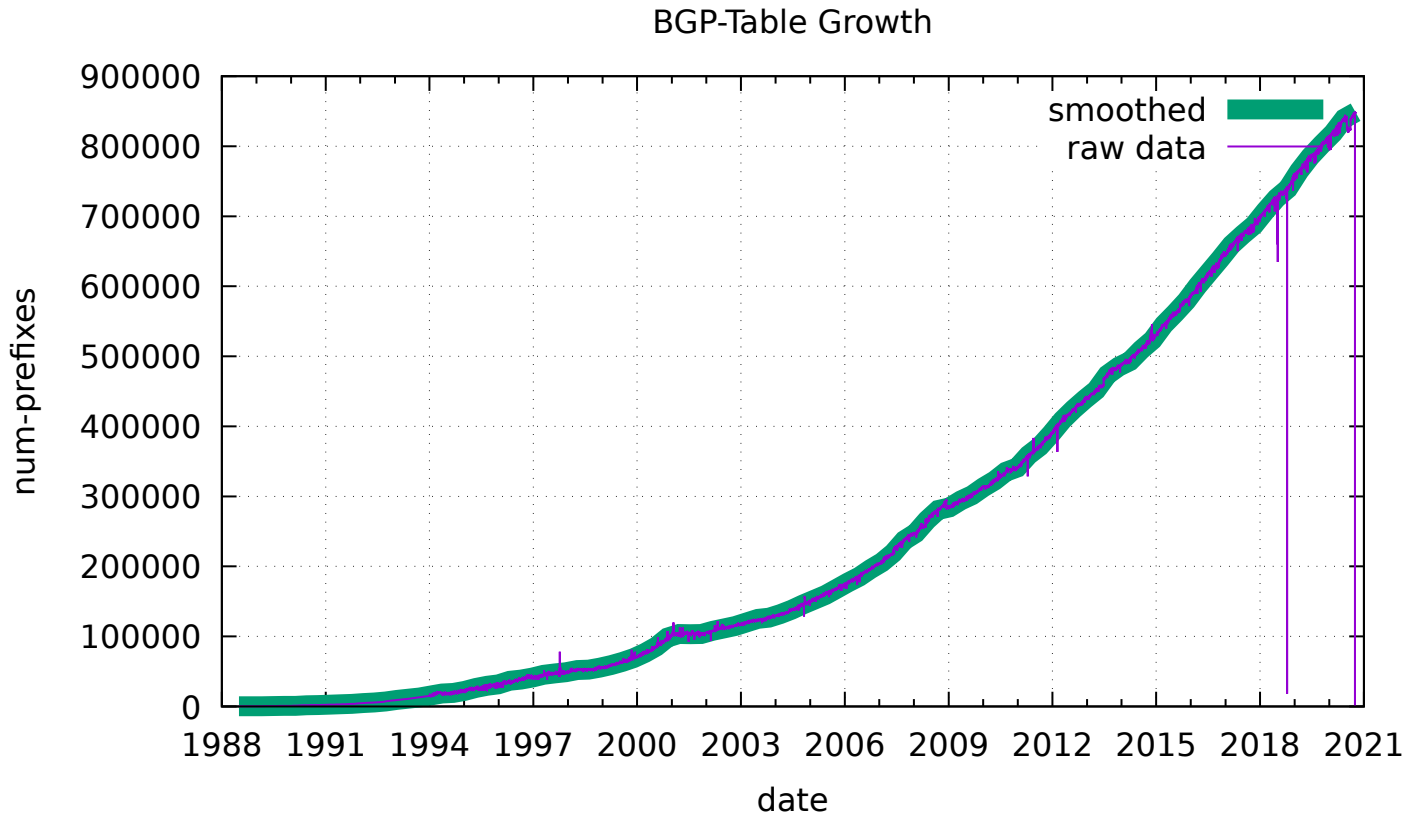
- Neustrukturierung[32] des Adressbereichs
- Klassen werden *nicht* mehr beachtet
- Bildung von *Supernetzen* für effizientes hierarchies routing[33]
  $\rightarrow$ kleine Routing-Tables
- Bessere Ausnutzung des Adressbereichs durch Aufteilung der A- und B-Klassen in kleinere Einheiten
- **neue Notation** als *Präfixlänge*[34]: 255.255.240.0$\rightarrow$/20

[32] soweit möglich... Bereits zugewiesene Netze können nicht entfernt werden

[33] eg. 212/7$\rightarrow$Europa

[34] "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 "<curl http://bgp.potaroo.net/as2.0/bgp-active.txt" using 1:2 title "bgp-table growth" 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

Reservierte/Spezielle IP-Prefix/Adressen

einige Bereiche im IP-Adressraum sind reserviert:

IP	Bedeutung	Source/Destination
0.0.0.0	unbekannte Source (DHCP/BOOTP)	nur Source
255.255.255.255	limited Broadcast	nur Destination, stoppt am Route
127.0.0.0/8	loopback, Host-lokal	beides, nur Host-intern
192.168.0.0/16	*private* IP, RFC1918, braucht NAT für Internet	beides, wird im Internet nicht geroutet
172.16.0.0/12	*private* IP, RFC1918, braucht NAT für Internet	beides, wird im Internet nicht geroutet
10.0.0.0/8	*private* IP, RFC1918, braucht NAT für Internet	beides, wird im Internet nicht geroutet
169.254.0.0/16	Link-Local, automatisch IPs ohne DHCP	beides, wird im Internet nicht geroutet
10.195.5.0/24	Netz-Basisadresse, alle Hostbits=0	keine
10.195.5.255/24	directed Broadcast, alle Hostbits=1	nur Destination

http://www.inetdaemon.com/tutorials/internet/ip/addresses/special.shtml

http://de.wikipedia.org/wiki/IP-Adresse#Besondere_IP-Adressen

http://www.rfc-editor.org/rfc/pdfrfc/rfc1918.txt.pdf

n|w

Routing und Routing-Table (1/3)

die Wegleitung – *routing* – im Internet:

- *Router* empfangen Pakete und leiten sie in Richtung Zieladresse[35] weiter: *routing* oder *forwarding*
- jedes Paket wird gesondert betrachtet[36]
- Pakete werden an den jeweils nächsten Router – *next-hop* – gesendet
- der "letzte Router"[37] stellt das Paket direkt an den Zielknoten[38] zu
- entschieden wird das alles über die *Routing-Tabelle*

[35] aus dem Paketinhalt/Layer-3 Adresse

[36] wichtig um die Robustheit des paketvermittelnden Netzes zu gewährleisten (moderne Router arbeiten möglicherweise effizienter)

[37] Router-Interface ist im selben Netzwerk wie die Zieladresse

[38] Endgerät, der Router findet die L2-Adresse mittels ARP

Routing und Routing-Table (2/3)

die *Routing-Tabelle* (RT) enthält (mindestens):

- *Ziel-Netz* ("dest-net" oder einfach "net")
- *Ziel-Maske* oder *Ziel-Präfixlänge* ("dest-mask", "prefixlength", "prefix")
- *Nächster-Router* ("next-hop" oder "gateway")

Beispiele:

```
Destination        Gateway            Flags        Refs      Use   Netif Expire
default            10.175.52.1        UGSc           12        0     en1
10.175.52/24       link#5             UCS             1        0     en1
10.175.52.1/32     link#5             UCS             1        0     en1
10.175.52.1        0:22:bd:f8:19:ff   UHLWIir        13       16     en1    869
```

```
Flags: X - disabled, A - active, D - dynamic, C - connect, S - static, r - rip, b - bgp, o - ospf, m - mme,
  B - blackhole, U - unreachable, P - prohibit
#      DST-ADDRESS        PREF-SRC        GATEWAY            DISTANCE
  0 ADS  0.0.0.0/0                         84.72.40.1                2
  2 ADC  84.72.40.0/21      84.72.42.48     ether1-gateway            0
  3 ADC  192.168.1.0/24     192.168.1.1     bridge1                   0
```

```
S*    0.0.0.0/0 [1/0] via 193.247.171.25
B     0.0.0.0/8 [200/0] via 192.168.1.1, 4w0d
      2.0.0.0/32 is subnetted, 1 subnets
B        2.58.228.145 [200/0] via 192.168.1.1, 2w3d
      5.0.0.0/32 is subnetted, 17 subnets
B        5.2.64.113 [200/0] via 192.168.1.1, 1d18h
B        5.2.64.133 [200/0] via 192.168.1.1, 1d19h
...
```

Routing und Routing-Table (3/3)

Ablauf "forwarding" – Paket weiterleiten

- finde passenden RT-Eintrag zur Zieladresse des weiterzuleitenden Pakets – für jede Zeile j der RT:
  target-ip $\wedge$ dest-mask$_j$ = dest-net$_j$
- falls ein passender Eintrag gefunden wurde, wird das Paket als Frame zur L2-Adresse der L3-Adresse `next-hop`$_j$ weitergeleitet
- falls der Router selbst eine IP-Adresse/Interface im Ziel*netz* hat, wird das Frame direkt an das Endgerät zugestellt
- die Routing-Tabelle wird nach absteigender[39] Präfixlänge abgearbeitet/sortiert
- die *Default-Route*[40] ist 0.0.0.0/0

[39] von "spezifisch" zu "allgemein"

[40] "passt" immer und wird am Schluss abgearbeitet

Intermezzo

Tools zum Thema

- `netstat -rn` zeigt lokale Routing-Tabelle an → finden des "Default-Router"
- `arp -a` zeigt die lokale ARP-Tabelle an → zur Kontrolle der Layer-2/MAC-Adresse des Routers
- `traceroute`[41] (unzuverlässiger) *Hin*weg[42] zu einer bestimmten L3-Zieladresse

[41] Windows: `tracert`

[42] zeigt die Router an, über die ein Paket wahrscheinlich weitergeleitet wird

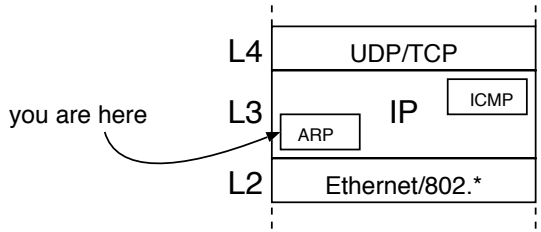
References

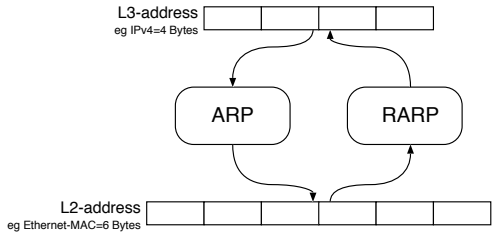
- http://en.wikipedia.org/wiki/IPv4_subnetting_reference
- http://en.wikipedia.org/wiki/Classless_Inter-Domain_Routing
- Route-servers: http://www.traceroute.org/#Route%20Servers
- BGP-routing-table growth: http://bgp.potaroo.net/as2.0/bgp-active.html, andere lustige Informationen: http://bgp.potaroo.net/as2.0/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

ARP: Address Resolution Protocol, (1/2)

- ARP findet zu einer gewünschten IP-Adresse die entsprechende

  MAC-Adresse: L3?→L2





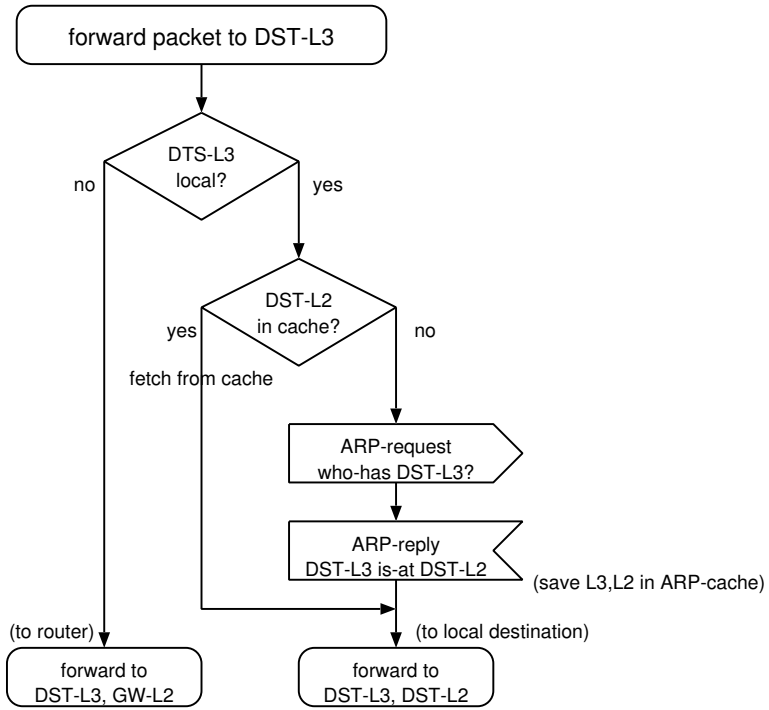
- RARP ist das "Rückwärtsprotokoll":
- ARP und RARP arbeiten beide mit L2/MAC-Broadcasts für die Anfragen und L2-Unicast für die Antworten[43]

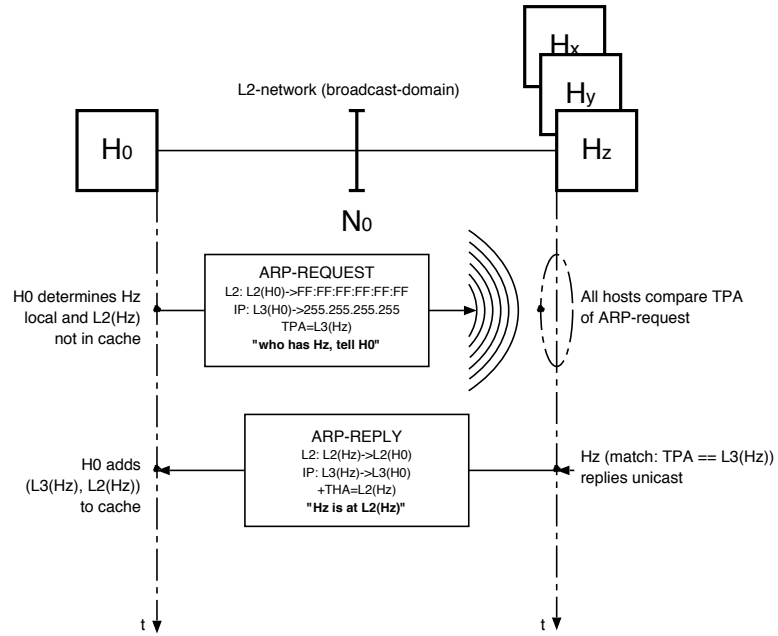
```
tcpdump -l -e -nnn arp
10:33:07.37 00:26:18:ce:27:6d > ff:ff:ff:ff:ff:ff, ethertype ARP (0x0806), length 42:
                                                   Request who-has 192.168.1.17 tell 192.168.1.16, length 28
10:33:07.37 00:0d:b9:16:bf:e4 > 00:26:18:ce:27:6d, ethertype ARP (0x0806), length 60:
                                                   Reply 192.168.1.17 is-at 00:0d:b9:16:bf:e4, length 46
```

[43] . . . meistens. Deshalb sehen Sie mit `tcpdump` nur die ARP-Anfragen (Broadcast)

ARP: Address Resolution Protocol, (2/2)

- Ablauf und Kommunikation

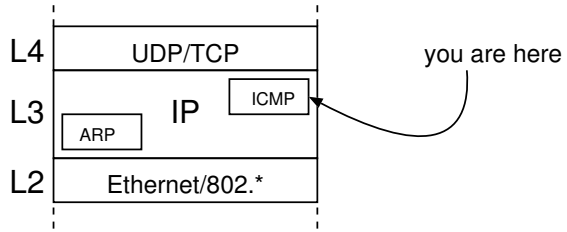




- Tools: ARP-Cache angucken, löschen oder statische Einträge einfügen: `arp`, `ip neighbor`

ICMP, Internet Control Message Protocol, (1/3)

IP (Layer-3) ist *best-effort*[44], d.h. es wird ein Mechanismus zur Fehlersignalisation benötigt:



- ICMP implementiert *Fehlermeldungen* und *Statusabfragen* in TCP/IP
- Router können durch ICMP Fehlerindikationen an das sendende Gerät zukommen lassen[45]
- Die Meldungen sind mit einem Code/Bedeutung markiert und enthalten den Original IP-Header des verursachenden Pakets[46]
- bei den meisten Fehlermeldungen wird das Original-Paket verworfen, der Sender muss versuchen es erneut zuzustellen oder die Fehlermeldung an die Applikation weiterzuleiten

[44] Pakete können verloren gehen

[45] . . . normalerweise sind Router "stumm", resp. dürfen nicht in die Kommunikation eingreifen

[46] um dem Gerät die Zuweisung des Fehlers an das betroffene Programm zu ermöglichen (z.B. browser)

ICMP, Internet Control Message Protocol, (2/3)

- Fehlermeldungen: ermöglicht Router und Endgeräte die Paketquelle über Fehler zu informieren (Auswahl)

Meldung	**Bedeutung**	**Sender**	Verworfen
network unreachable	kein passender Routing-Table Eintrag	Router	Paket verworfen
host unreachable	keine Antwort auf ARP	letzter Router	Paket verworfen
port unreachable	kein Serverprozess, Listen-Socket	Zielhost	Paket verworfen
time exceeded	TTL abgelaufen[47]	Router	Paket verworfen
fragmentation needed	Paket zu gross	Router	Paket verworfen
redirect	an anderen Router senden[48]	Router	Paket weitergeleitet
source-quench	Flusskontrolle, veraltet[49]	Router	Paket weitergeleitet

- Statusabfragen: ermöglicht einfache Statusabfragen auf Layer-3

Meldung	**Bedeutung**	**Sender**
echo-request und echo-reply	Erreichbarkeitstest[50] auf Layer-3	alle
timestamp-request und -reply	Zeitstempel[51] Abfrage	alle

[47] *Time To Live* ist ein numerisches Feld im IP-Header und wird von jedem Router dekrementiert (-1) – bei 0 → time-exceeded

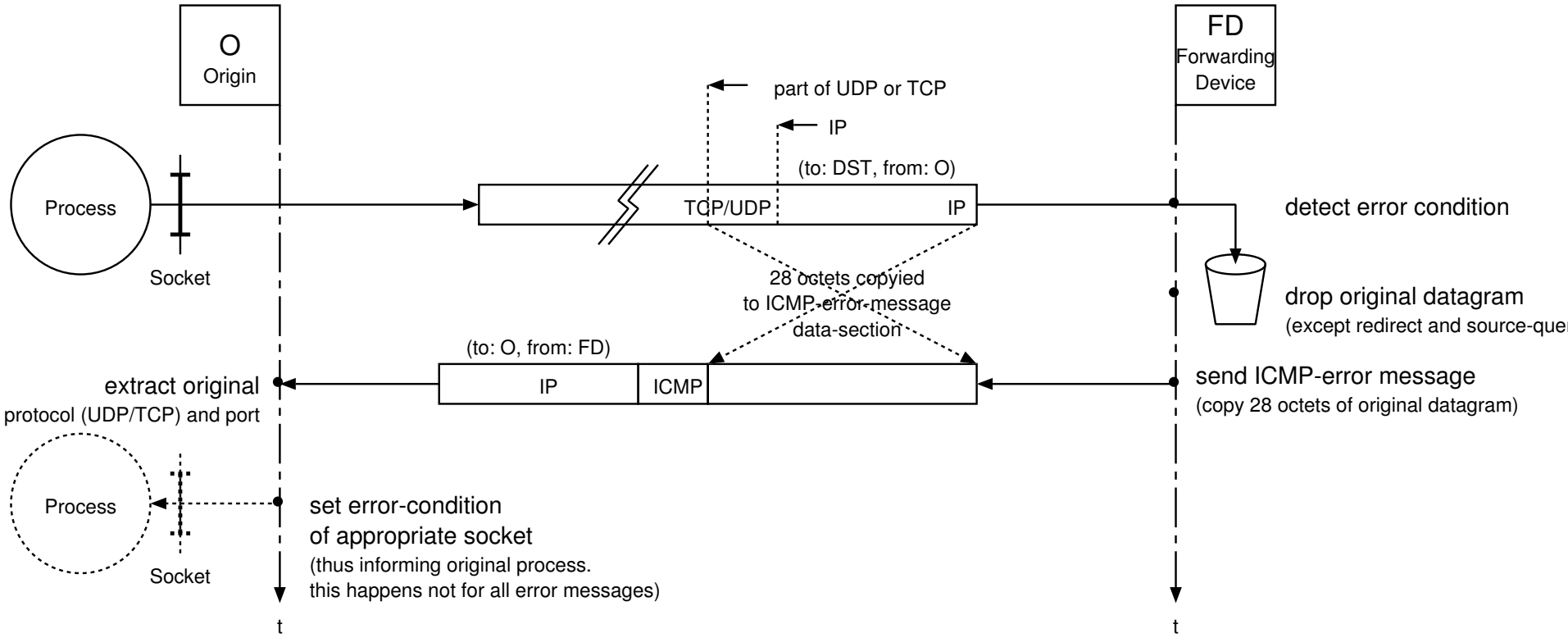
[48] enthält IP des "besseren" Routers → Routing-Table

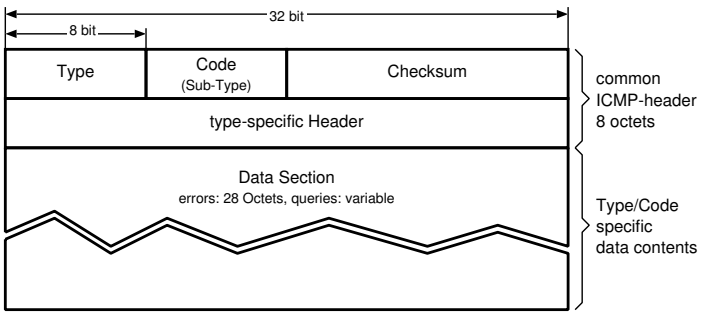
[49] . . . da meistens ignoriert und zusätzlicher Aufwand (Status) für Router

[50] z.B. `ping`

[51] enthält Local-Absende-, Remote-Empfangs- und Absende-Zeitstempel. Local-Empfangszeitstempel wird bei Erhalt der -reply-Meldung eingetragen

ICMP, Internet Control Message Protocol, (3/3)





ICMP Tools (1/2)

- `ping`: Layer-3 reachability testen Erreichbarkeit[52] auf Layer-3:

```
rschmutz@callisto ~ $ ping -c 3 www.google.ch
PING www.l.google.com (209.85.227.99): 56 data bytes
64 bytes from 209.85.227.99: icmp_seq=0 ttl=54 time=44.935 ms
64 bytes from 209.85.227.99: icmp_seq=1 ttl=54 time=41.854 ms
64 bytes from 209.85.227.99: icmp_seq=2 ttl=54 time=58.240 ms

--- www.l.google.com ping statistics ---
3 packets transmitted, 3 packets received, 0.0% packet loss
round-trip min/avg/max/stddev = 41.854/48.343/58.240/7.110 ms
```

- `ping`/ICMP-Echo-Request testet nur die *Erreichbarkeit auf Schicht-3 "End-zu-end System"*[53]
- ICMP-Echo-Request/Response wird häufig auf Firewalls "geblockt"
  - `ping` auf eine Adressen/Host ergibt keine Antwort (timeouts) *aber* der HTTP-Dienst auf derselben Adresse/Host ist verfügbar[54]
- die Ausgabe ist normalerweise die "Round-Trip-Time", d.h. die total benötigte Zeit Anfrage+Verarbeitung+Antwort

[52] mit Hilfe von ICMP-echo-request Meldungen

[53] d.h. IP-Adresse/Host kann erreicht werden aber keine Angaben über die Verfügbarkeit eines speziellen Dienstes (Mail, Web, etc) auf diesem Host

[54] z.B. `ping www.microsoft.com`→nix aber `http://www.microsoft.com/` im Browser funktioniert

ICMP Tools (2/2)

- `traceroute`: Layer-3 route path[55]

```
rschmutz@zaphod:~$ traceroute www.google.com
traceroute to www.google.com (209.85.135.103), 30 hops max, 40 byte packets
 1  static.193.65.40.188.clients.your-server.de (188.40.65.193)  0.676 ms  0.702 ms  0.723 ms
 2  hos-tr1.juniper1.rz10.hetzner.de (213.239.227.129)  0.189 ms  0.194 ms
    hos-tr4.juniper2 (213.239.227.225)  0.178 ms
 3  hos-bb1.juniper2.ffm.hetzner.de (213.239.240.226)  4.559 ms  4.576 ms  4.597 ms
 4  de-cix10.net.google.com (80.81.192.108)  5.682 ms  5.981 ms  6.331 ms
 5  209.85.255.172 (209.85.255.172)  6.098 ms
    209.85.255.170 (209.85.255.170)  16.070 ms  6.192 ms
 6  72.14.238.128 (72.14.238.128)  14.460 ms  14.001 ms
    209.85.248.248 (209.85.248.248)  11.706 ms
 7  209.85.241.187 (209.85.241.187)  13.981 ms
    209.85.241.83 (209.85.241.83)  13.886 ms  14.033 ms
 8  209.85.253.22 (209.85.253.22)  13.074 ms  13.896 ms
    72.14.239.54 (72.14.239.54)  27.321 ms
 9  mu-in-f103.1e100.net (209.85.135.103)  15.272 ms  15.526 ms  13.659 ms
```

- `traceroute` zeigt den scheinbaren "Pfad"[56] – d.h. die einzelnen Router + Distanz in "hops" – zu einer Zieladresse
- dazu werden TTL-begrenzte Pakete[57] ausgesendet und die ICMP-Time-Exceeded-in-Transit Meldungen der Router ausgegeben
- wie auch `ping` wird `traceroute` häufig "geblockt" und zudem sind die Angaben interpretationsbedürftig

[55] durch erzwingen von ICMP-time-exceeded und ICMP-port-unreachable Meldungen

[56] IP=Paketvermittelnd → *kein* "Pfad", d.h. `traceroute` lügt ein bisschen

[57] z.B. UDP-Pakete an möglichst unbenutzte Ports