

Network Layer
Ende zu Ende Verbindung zwischen Sender u. Empfänger
Router interessieren sich nicht für Schicht 4 / 5. Nur Weiterleitung zum Zielhost
Adressierung: IP Adressen
Forwarding: Welches Ausgangsinterface des Routers. Bei Router oft in HW implementiert
Routing: Berechnung der Wege im Netz. Eintragen d. Ergebnisse in **Weiterleitungstabellen** (IP Adr. Reichweiten jeweils). Routingprotokolle (konstruieren Routing Tabellen). Meist in SW implementiert.
IP ist verbindungslos.
Link Layer kann unterschiedlich sein (WLAN, Ether...)
Best Effort: Jeder Router tut sein Bestes, aber keine Garantie bzgl Reihenfolge, Bandbreite... und keine Bevorzugung

IP Adresse
32 Bit. Identifiziert Host im Internet, gehört aber logisch eigentlich zum **Interface des Hosts**. **Jedes Interface (am Router mehrere) hat eigene IP-Adresse.**
Subnetze
Mehrere Hosts teilen sich gleichen IP Adresspräfix. Innerhalb von Subnetz können sich Hosts ohne Router erreichen.
Vorteil: In Routingtabellen müssen nur Subnetzadressen stehen
Beispiel: 223.1.3.0/24 (die ersten 24 Bits sind für alle Hosts des Subnetzes gleich)

ARP (Address Resolution Protocol)
Herausfinden, welche Ziel-MAC zum Next-Hop Router/Host gehört. ARP Paket wird in Ethernet Frame verpackt
→ Übersetzen von IP in MAC Adressen
ARP Tabelle in allen Hosts (nicht Switches)

Arp Tabelle: IP | MAC
Routing Tabelle: IP | Ausgangsport

Switch forwarding Tabelle: MAC | Ausgangsport
- ARP Paket direkt in Ethernet Frame verpackt => kein IP Header

Links Seite vs Distance Vector
Routingnachrichten:
LS: Jeder Router flutet Infos über seine Links im ganzen Netz DV: Jeder Router informiert seine Nachbarn welche Ziele er zu welchen Kosten erreichen kann

Robustheit: Was wenn ein Router bösartig ist?
B's Router kann falsche Linkkosten ankündigen. Fehler begrenzt, da jeder Router seine eigene Tabelle berechnet.
DV: Router kann fiasche Pfadkosten ankündigen. Fehler plizen sich fort, da Tabelle eines Routers Einfluss auf andere Router hat.

BGP
Policy-based Routing: Beispiele:
- Ignoriere Pfade durch AS Y
- Gib Routinginfo nicht an Nachbarn AS X weiter
BGP, OSPF: Welches Ausgangsinterface muss Router verwenden, um zum Gateway Router zu kommen, der über iBGP mitgeteilt wird.
Beste Route nach Kriterien:
1: Link Pref, 2: Kürzester AS Pfad, 3: Route mit am schnellsten erreichbarem Next-Hop (Hot Potato)..
Hot Potato Routing
AS will, dass Pakete so schnell wie möglich das eigene AS verlassen, also lokales Gateway mit **geringsten Intradomain Kosten** wählen.

Pipelining
Sender darf mehrere Paket gleichzeitig senden, aber es dürfen nur eine begrenzte Anzahl von unbestätigten Paketen unterwegs sein.

Go-Back-N
- Empfänger bestätigt immer nur mit Sequenznummer für die gilt, dass alle kleineren bereits empfangen.
- Sender hat Timer für ältestes unbest. Paket. Bei Timeout wrden **alle** Pakete nochmal geschickt, die noch nicht bestätigt wurden.

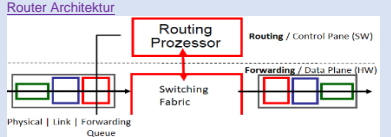
Selective Repeat
- Retransmission nur für verlorene Pakete.
- Empfänger schickt ACKs für jedes einzelne Paket individuell.
- Sender hat Timer für jedes Paket. Timerablauf -> nur betreffendes Paket wird neu gesendet.

Hybrid von Go-Back-N und Selective Repeat
Von Go-Back-N:
- Kummulative ACKs
- Nur 1 Retransmission Timer. ältest unbestät Segment
Von Selective-Repeat:
- Empfangspuffer
- Bei Timeout wird nur das verlorenegegangene Paket erneut gesendet
Auslösen von Retransmissions durch:
- Timeouts
- Duplikat ACKs (3) => **Fast Retransmit:** 3x Duplicate ACK: Indiz für Paketverlust
Sender erhält **mehr als 3x** gleiches ACK
=> Retransmission des ältesten unbestätigten Segment (Seq mit der ACK Nummer)

Flow Control (Empfänger zu langsam)
Freier Puffer **rwnd** = RcvBuffer - (LastByteRcvd - LastByteRead)
Empfänger teilt Sender rwnd mit.
Congestion Control (Netzwerk zu langsam)
2 Ansätze:
Netzwerk-unterstützt: Router geben Rückmeldung an Hosts bei Überlastung.
Ende-zu-Ende: Auslastung des Netzwerks wird durch Beobachten der Verzögerungen und Auftreten von Paketverlusten abgeschätzt. (TCP!!)
Additive Increase: Nach jeder RTT wird cwnd um 1 MSS erhöht bis Paketverlust erkannt
Multiplicative Decrease: Halbiere cwnd nach erkanntem Paketverlust

Window Size <= min(cwnd,rwnd)

TODO: letztes Thema



Queueing, falls Ankunftsrate schneller als Weiterleitung durch Fabric. Link Layer wird hier terminiert.

Bei IP wird Ausgangsportal nur anhand der IP Zieladresse bestimmt!
Longest Prefix Matching um Platz in Tabelle zu sparen!

Classful Addressing
Früher: **Feste Länge** für Subnetzpräfixe (/8, /16, /24) /24 Netz kann 2^N(32-24) = 2⁸ Hosts haben
Classless Addressing
Beliebige Länge für Subnetzpräfixe (CIDR)
Präfixnotation: 200.23.16.0/24 zeigt, welche Bits zum Subnetz gehören
Spezielle IPv4 Adressen
127.0.0.1: Localhost, eigener PC. Netmask 255.0.0.0
Private IPv4 Adressen: 10.0.0.0/8, 172.16.0.0/12, 192.168.0.0/16, 169.254.0.0/16

Beispiel: 192.168.0.0/16 Netzmaske: 255.255.0.0
Broadcast: 192.168.255.255 Netzadresse: 192.168.0.0

ARP Ablauf (Sender und Empfänger in einem Netz)
A möchte Datagramm zu B senden
B's MAC nicht in A's ARP Tabelle

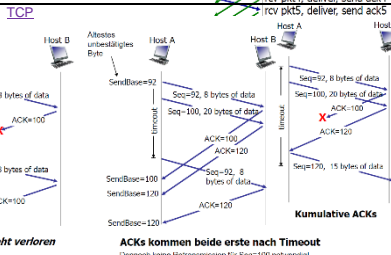
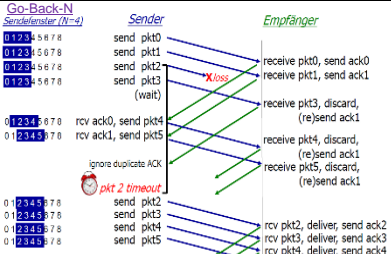
A schickt Broadcast ARP Query Paket das B's IP Adresse enthält (Ziel MAC: FF-FF-FF-FF-FF-FF)
Alle Hosts im LAN empfangen diese ARP Query

B empfängt ARP Query und informiert A in Antwort über B's MAC Adresse -> Unicast Frame zu MAC A

A speichert IP/MAC paar in ARP Tabelle bis veraltet
ARP bedarf keiner Konfiguration (plug and play)

Hierarchisches Routing
Router werden in autonome Systeme (AS) gegliedert.
Intradomain Routing (für Ziele im gleichen AS)
Wie sind Gateways zu Nachbarbarzeten aus lokalem Netz erreichbar?
Protokolle: **RIP, OSPF, IGRP**
Interdomain Routing (für Ziele in anderen ASen)
Welche externen Ziele sind über welches Transfer-AS / Gateway erreichbar. (Gateway=Router am Rande des next AS)
Protokolle: **BGP**
Zusätzliche Routing Policies notwendig:
- Jeder Router kann bestimmen, welche Nachrichten er bevorzugt.
- Auch wirtschaftliche Aspekte spielen eine Rolle.

IPv6
40 Byte Header Kein ARP
Keine Fragmentierung und keine Checksumme
Notation: 128 Bit in 8 Blöcke, je 16 Bit (4-6 Hex Zahlen), mit „::“ getrennt.
- Führende Nullen darf man weglassen
- EINMAL dürfen ein oder mehr aufeinanderfolgende Blöcke mit 0000 ausgelassen werden und durch „::“ ersetzt werden
- Beispiel: 2001:0db8:0:0:0:0:1428:57ab wird zu: 2001:db8::1428:57ab Nullen am Ende: 2001::32:
- Es gibt keine /80 Subnetze, weil Host-ID immer genau 64 Bit.
- IPv4 Notation: <192.31.20.46
- **Tunneling:** IPv6 Paket in IPv4 Paket (bei legacy Leitung)
Link Local Adresse: im Bereich fe80::10 (Für Hosts im gleichen Subnetz) wird direkt aus MAC Adresse abgeleitet



TCP Reno
Slow Start: Verdopple cwnd nach jeder RTT bis ssthresh
Congestion Avoidance: Dann - vergrößere cwnd um 1 nach jeder RTT
Router Timeout: TCP Slow Start | ssthresh = cwnd/2 | cwnd = 1
Nach 3 Duplicate ACKs: Congestion Avoidance, ssthresh = cwnd/2, cwnd = cwnd/2 + 3
Diagram: Shows packet loss, timeout, and recovery process with dup ACKs and ssthresh adjustments.

TODO: letztes Thema

Switching Fabric
3 Typen: Memory, Bus, Crossbar
Queueing an Eingangsports
Nötig, falls Fabric langsamer als Ankunftsrate.
Head-of-the-Line Blocking: Vorderstes Paket blockiert andere Pakete, obwohl andere Pakete zu einem Ausgangsinterface müssen, das frei ist.
Queueing an Ausgangsports
Nötig, falls Ankunftsrate von Fabric die Übertragungsgeschwindigkeit des Ausgangslinks übersteigt.
Wenn Queue voll -> **Paketverluste!**

DHCP (Dynamic Host Configuration Protocol)
Server weist automatisch IP Adressen zu.
Manuelle zuweisung auch möglich (siehe Win/Linux Befehle)
Eigentlich Schicht 4!
Host kann zugewiesene IP Adresse ggfs verlängern.
Ablauf:
- Host sucht DHCP Server: DHCP Discover (optional)
Ziel: 255.255.255.255 (Ethernet Broadcast)
- DHCP Server antwortet mit DHCP Offer (optional)
Ziel: 255.255.255.255 (Ethernet Broadcast)
- Host fordert explizit IP Adresse an: DHCP Request
Ziel: 255.255.255.255
- DHCP Server weist Adresse zu: DHCP ACK
Ziel: 255.255.255.255

ARP Ablauf (Sender u. Empfänger in versch. LANs)
Annahme1: A kennt IP Adr von B über DNS
Annahme2: A kennt IP und MAC von Router R

- A erzeugt IP Datagramm mit SourceIP A und DestIP B
- A erzeugt Link-Layer Frame mit R's MAC Adresse
- Frame wird von A nach B geschickt

- R empfängt Frame, entfernt Ethernet Header und gibt Inhalt hoch zu Network Layer
- R leitet IP Datagramm mit SourceIP A und DestIP B weiter
- R erzeugt Ethernet Frame mit B's MAC als Ziel, Frame enthält IP Paket von A zu B

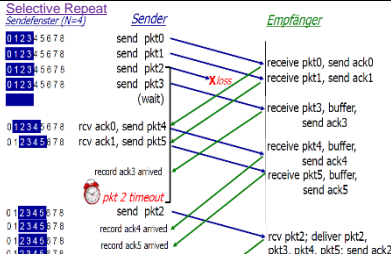
OSPF (Open Shortest Path First)
Router fluten **Link State** Advertisement Nachrichten (enthalten Infos über alle Nachbarrouter) an alle anderen Router im gesamten AS. Dann Routenberechnung über Dijkstra.
OSPF Advertisements werden direkt über IP gesendet (kein TCP oder UDP).

Router muss lernen, über welches Interface komme ich aus meinem AS raus zum Ziel, welches sich evtl. in einem ganz anderen AS befindet*. Das steht dann in der **Routing Tabelle** der Router

Transport Layer
Kommunikation zwischen **Prozessen** auf Sender u. Empfänger Seite. (Betrifft Hosts, nicht Router!)
Teile des **Betriebssystems**

Transport Layer Multiplexing (in TCP und UDP)
IP Pakete werden Prozessen des BS zugeordnet.
Port: 16 Bit.

UDP
Verbindungslos: Socket definiert durch: **Dst IP/Port**
UDP und TCP, garantieren nicht Delay/Bandbreite!
Übertragungsfehler über Checksumme erkennbar.



TCP Verbindungsaufbau (3-Way-Handshake)
Verbindungsaufbau = Overhead
Wähle initiale Seq. n; sende TCP SYN Nachricht
SYNbit=1, Seq=x
Wähle initiale Seq. n'; sende TCP SYNACK Nachricht; bestätige SYN
SYNbit=1, Seq=y
ACKbit=1; ACKnum=x+1
Empfang von SYNACK(y) ist Hinweis, dass Server lebt; sende ACK für SYNACK; hier können bereits Nutzdaten enthalten sein
ACKbit=1, ACKnum=y+1
Empfangenes ACK(y) ist Hinweis Client lebt

NAT (Network Address Translation)
Lokales Netzwerk benutzt nur 1 IP Adr. um mit Rest des Internets zu kommunizieren.
16 Bit Portnummer-Feld => >60000 gleichzeitige Verbindungen mit 1 öffentlichen IP.
Umstritten, da Schicht 4 (Portnummer) nicht vom Router berücksichtigt werden sollte.
Port forwarding:
Verbindungsanfragen an bestimmten Port werden fest zu bestimmten Server weitergeleitet.
Beispiel: (123.76.29.7, Port 2500) wird immer zu (10.0.0.1, Port 25000) weitergeleitet.
Manuelle Konfiguration notwendig!

NAT Tabelle:

TODO: letztes Thema

IP (20 Bytes Overhead für IP Header)
- Adressierungskonventionen TTL: Anz Hops
- Datagram Format verbleibend
- Packet handling conventions Nutzlast=size-20Byte
IP Fragmentierung

MTU: Verschiedene Link Layer Technologien haben versch. max. Paketgrößen (Ethernet 1500 Byte)
Router/Host zerlegt in kleinere Pakete.
Zusammenbau am End-Host! (Betriebssystem Overhead)
16-Bit Identifier: Identisch für alle Pakete eines Frames
Fragmentation Flag: 1: Da kommt noch was, 0: letztes Fragment eines Pakets
Offset: Byteposition innerhalb des Gesamtpaketes, an die das Fragment gehört. (Offset 185: 185 * 8 = 1480, also nach Byte 1480 kommt dieses Fragment)

ICMP (Internet Control Message Protocol)
Error reporting und Router signaling.
Austausch von Infos zw. Host und Routern.
- ICMP Information wird als IP Paket versendet

Type	Code	description
0	0	echo reply (ping)
3	0	dest. network unreachable
3	1	dest host unreachable
3	2	dest protocol unreachable
3	3	dest port unreachable
3	6	dest network unknown
3	7	dest host unknown
4	0	source quench (congestion control - not used)
8	0	echo request (ping)
9	0	route advertisement
10	0	router discovery
11	0	TTL expired
12	0	bad IP header

Inhalt ICMP Nachricht: Type, Code
Erste 8 Bytes des IP Pakets, das Fehler verursacht

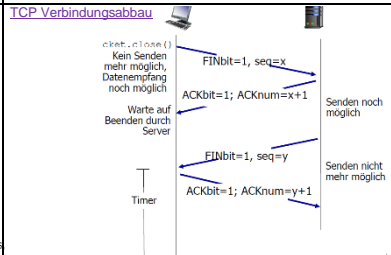
Routing
Router tauschen Kontrollnachrichten aus, um Routingtabelle zu erzeugen.

Link State (zentral): Jeder Router kennt komplette Topologie. Z.B. OSPF
Distance Vector (dezentral): Jeder Router kennt nur direkten Nachbarn u. Kosten zu diesen. Nachbarn teilen per Routingnachrichten mit welchen Knoten sie mit welchem Gesamtkosten erreichen können. Z.B. Routing Information Protocol (RIP).
Statisch: Manuelle Konfiguration von Forwardingtable
Dynamisch: Periodischer Austausch von Routinginformationen. Änderungen autom. erkannt.

BGP (Border Gateway Protocol)
Teile Dem Rest der Welt die Existenz eines IP Präfix mit.
eBGP: Verbindungen **zwischen** ASen. Teilt anderen ASen Mit, dass man über diesen Router zu einem bestimmten IP Subnetz kommt. (TCP Sessions zwischen Gateways in ASen)
iBGP: Verbindungen **innerhalb** eines AS. Teilt anderen Routern innerhalb eines AS mit, dass man über diesen Router zu einem bestimmten Ziel kommt. (TCP Sessions zwischen allen Routern innerhalb eines AS)
BGP Attribute:
- AS-PATH: Liste von ASen, durch die Prefix Advertisement gehlaufen ist. (Von aktuellem bis Zielnetz)
- NEXT-HOP: Router im eigenen AS, der das AS mit dem nächsten AS auf dem AS Pfad verbindet.

TCP (alles in SW implementiert)
Verbindungsorientiert: Socket definiert durch:
Src IP/Port, Dst IP/PORT Src Port meist beliebig wählbar
Zuverlässig: Keine Bitfehler, kein Datenverlust, korrekte Reihenfolge
Problem: Zuverlässige Übertragung über unzuverlässigen Kanal (IP)
ACK: Um zu sagen, ob die Daten angekommen sind, und OK (Checksum) sind.
Sequenznummern: Um zu sehen, ob ACK Corrupt ist und um zu sehen, ob das letzte Paket richtig angekommen ist (Ersatz für NACK).
Timeout: Schützt vor Daten/PACKs die auf dem Weg verloren gehen.
Pipelining: Für verbesserte Performance. (nicht mehr Stp n Wt)

TCP
Zuverlässig: Bestätigt einzelne Bytes, nicht Pakete!
Pipelining: Mischung aus Go-Back-N u. Selective Repeat.
Flow Control u. Congestion Control bestimmen Fenstergr.
Vollduplex: Max Segment Size (MSS) richtet sich nach MTU der Link Layer
Verbindungsorientiert: Verbindungsaufbau vorher. Sender u. Empfänger initialisieren State Maschine



Alle Datagramme, die das lokal Netz verlassen, haben **gleiche** Src IP, aber versch. Src Ports

WAN Seite/Internet	IP Adr	Port	LAN Seite/Heim-Netz	IP Adr	Port
239.120.38.24	5100	192.168.0.4	9929		
239.120.38.24	5101	192.168.0.5	9929		
239.120.38.24	5102	192.168.0.6	9929		
239.120.38.24	5103	192.168.0.4	9930		
...		

TODO: letztes Thema

Abkürzungen

Einführung

DSL: Digital Subscriber Line
ISP: Internet Service Provider
TCP: Transmission Control Protocol (Netzwerkprotokoll, das definiert auf welche Art und Weise Daten zwischen Netzwerkkomponenten ausgetauscht werden sollen)
UDP:
IP: Internet Protocol (Protokoll das die Grundlage des Internets darstellt)
HTTP: Hypertext Transfer Protocol (Protokoll zur Übertragung von Daten in der Anwendungsschicht)
RFC: Request for Comments (legt Internet Standards fest)
VoIP: Voice over IP
DSLAM: Digital Subscriber Line Access Multiplexer (übersetzt hochfrequente Töne in digitale Signale, bevor Daten zum Modem im Heimnetz kommen)
CMTS: Cable Modem Termination System (Wie DSLAM aber für Kabelmodem)
DHCP Server: Dynamic Host Configuration Protocol Server (Verteilt automatisch Adressen an Hosts in einem Netzwerk)
DNS Server: Domain Name System Server (Weist im Internet einer URL die richtigen IP-Adresse zu)
SAP: Service Access Point (Im Schichtenmodell stellt jede niedrigere Schicht der jeweils höheren Schicht einen SAP zur Verfügung. Somit kann die Höhere Schicht die Services der niedrigeren benutzen)
ISO: International Organization for Standardization
OSI: Open Systems Interconnection

Network Layer

CIDR: Classless Interdomain Routing (Subnetzteil einer Adresse kann beliebige Länge haben)

Windows/Linux Befehle

Messen der Round Trip Time:

- Windows: ping, mehrere Pings: ping -a 10000

Wege eines Pakets durch das Internet verfolgen:

- Windows: tracert
- Linux: traceroute

Routing Tabelle anzeigen:

- Linux: route
- Windows: route print

Route in Routingtabelle hinzufügen:

- Linux: ip route add 100.0.2.0/24 via 100.0.1.2

Forwarding aktivieren:

- Linux: sysctl -w net.ipv4.ip_forward=1

TCP Server starten:

- Linux: netcat -l -p 9000

Mit TCP Server verbinden:

- Linux: netcat <dstIP> 9000

MAC-Adresse rausfinden:

- Windows: ipconfig /all -> physische Adresse
- Linux: ifconfig -> ether, oder ip addr

MAC-Adresse ändern:

- Windows: Systemsteuerung -> Geräteanalyzer

Adresszuweisung

- Linux: ifconfig eth0 200.23.16.4 netmask 255.255.255.0 **oder** ip addr add 200.23.16.4/24 dev eth0 **oder** persistent: /etc/network/interfaces
- Windows: Systemsteuerung -> Netzwerk- und Freigabecenter -> Adaptereinstellungen

IP Adresse von DHCP Server anfordern

- Linux: dhclient
- Windows (ipconfig /release)

ARP-Tabellen anzeigen

- Linux: arp
- Windows: arp -a

Eintrag aus ARP-Tabelle löschen

- Linux: arp -d <IP-Adresse>

nmap – Portscanner, scannt auf offene Ports im Netzwerk (TCP SYN Scan). Falls ein Port offen ist, wird ein TCP SYN ACK vom Zielhost an Scanner zurückgeschickt. Benutzt ARP Requests zur Erkennung von aktiven Hosts

TCP Verbindungen anzeigen:

- Windows: netstat -p tcp

WIRESHARK

Filter für MAC destination:

- z.B. Broadcasts finden: **eth.dest == ff:ff:ff:ff:ff:ff**

TODO

Evtl. Übung 7 1.

Was ist ein IXP