

Ethernet: Prinzipien

- 10 Mb/s; Manchester-Kodierung (20 MHz; +/- 0.85V)
- Nachricht wird in Pakete ("frames") fester Länge zerteilt
 - bei einem Übertragungsproblem lediglich (kurzes) Paket wiederholen
 - Paketbildung, Bitfehlerprüfung durch die Netzadapterkarte

- Sendewillige Station horcht am Bus ("Carrier Sense") und sendet, falls Medium frei

- ansonsten warten bis frei und dann sofort senden
- Problem: Tun das dann nicht oft mehrere Stationen gleichzeitig?
- nach Senden eines Paketes einige μ s warten (Lücke ist Chance für andere sendewillige Stationen)

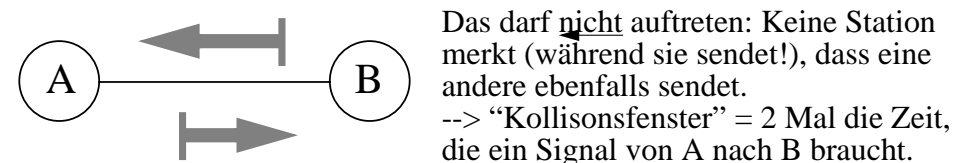
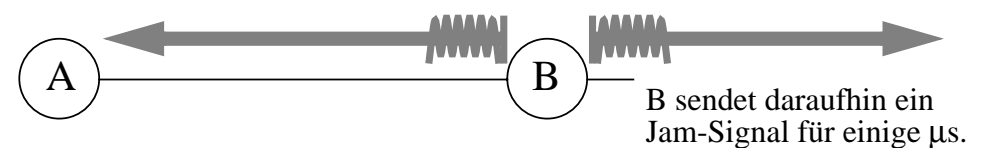
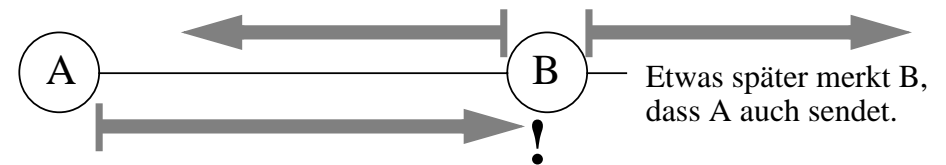
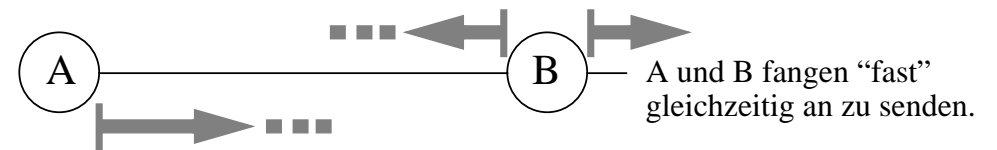
- Sendende Station horcht während des Sendens Bus ab; Kollisionen mit anderen Stationen werden vom zugehörigen Transceiver erkannt ("Collision Detection")

- endliche Ausbreitungsgeschwindigkeit --> minimale Paketlänge
- bei Kollision: Senden abbrechen

- Bei Kollision neuer Sendeversuch nach zufällig bestimmter Zeit aus einem Intervall

- bei jeder erfolglosen Wiederholung wird Intervall verdoppelt ("binary exponential backoff"); initial: 51.2 μ s
- Abbruch nach 16 Fehlversuchen (Paket hätte ja auch sonstwie verloren gehen können --> Behandlung auf höherer Ebene)

Kollisionen

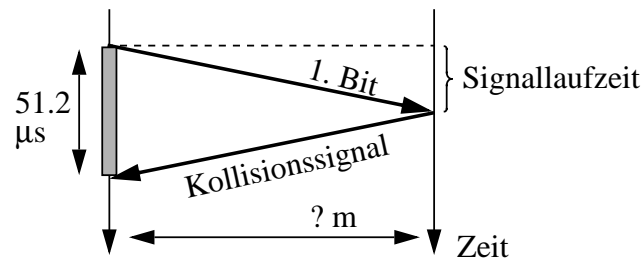


- Kollision wird erkannt:

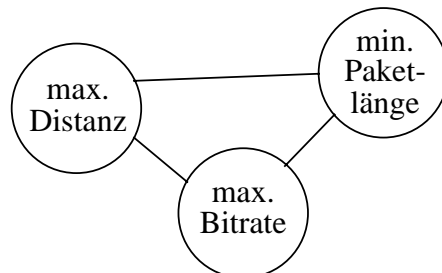
- bei ungewöhnlich hoher Spannung, erzeugt durch mehr als einen Sender
- bei falscher Impulsbreite
- ggf. auch durch Vergleich gesendete / selbst wieder empfangene Daten

Kollisionsfenster

- Mindestpaketlänge: 64 Byte
 - kürzere Nachricht (z.B. bei Abbruch eines Sendevorgangs) werden künstlich verlängert
- Sendevorgang dauert also mindestens 51.2 μ s
- Zur sicheren Kollisionserkennung muss Kollisionssignal ankommen, bevor das letzte Bit des Paketes versendet wird

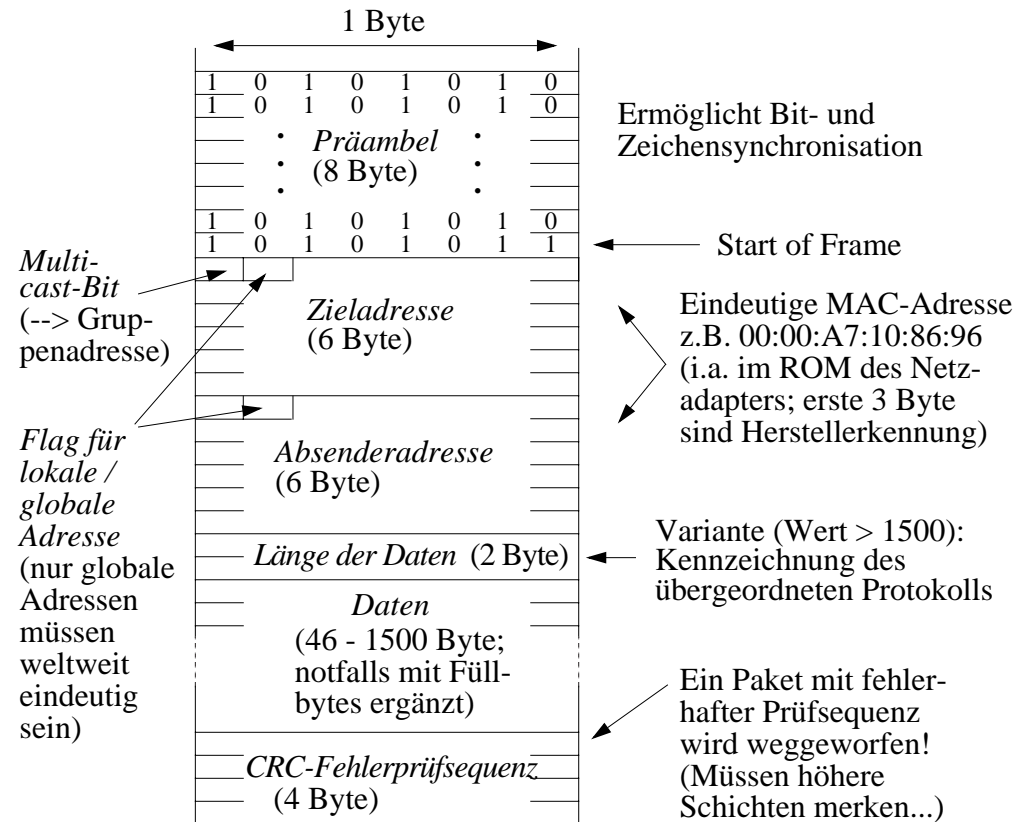


- Bei Signalgeschwindigkeit $> 0.6c$ --> ca. 4500 m, die zwei Stationen maximal auseinanderliegen dürfen
 - tatsächlich laut Standard: 2500 m
 - Repeater verzögern geringfügig (max. 8 Bitzeiten = einige 100 m)
- Es ergibt sich damit eine Anhängigkeit von drei Größen, die bzgl. des Designs einen Zielkonflikt darstellen:



- Frage: Ist das Prinzip für Gb/s-Netze noch praktikabel?

Ethernet - Paketformat



- Minimale Paketlänge = 64 Byte (ohne Präambel), damit Kollisionen sicher erkannt werden
 - wenn eine gewisse Zeit (= Mindestzahl gesendeter Bits!) seit Beginn der Sendung vergangen ist (und keine Kollision entdeckt wurde), kann der Sender wegen der Maximallänge des Netzes sicher sein, dass er das Medium exklusiv hat...
 - grosser Overhead bei einzelnen Zeichen (z.B. Tastatureingabe)
 - bei fehlendem Längenfeld wird Paketende an nachfolgender Sendepause erkannt (min. 9.6 μ s; entspricht 12 Byte)

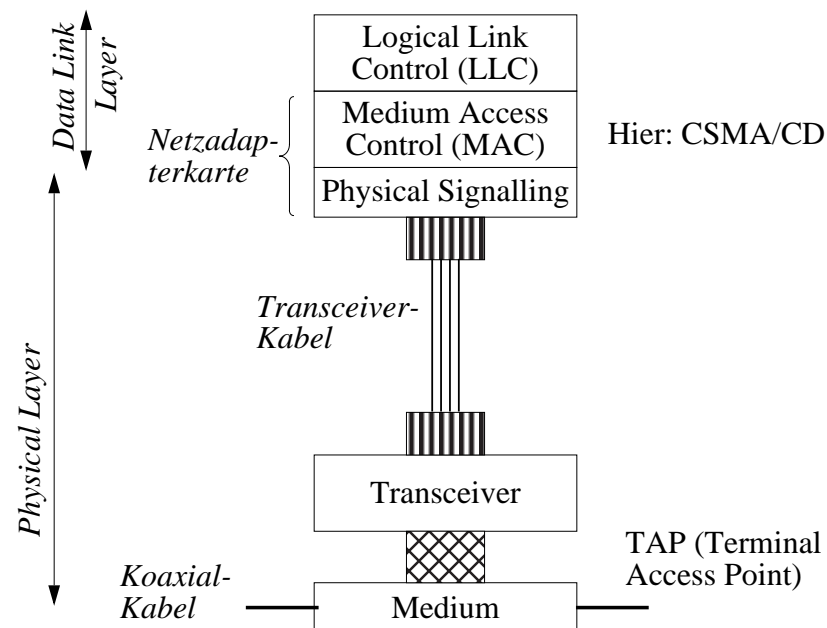
MAC-Adressen

- MAC (= Medium Access Control) ist die untere Hälfte der Schicht-2-Protokolle bei LANs
 - obere Hälfte: LLC (Logical Link Control)
- MAC-Adresse: 6 Byte (= 48 Bit) lang
 - andere Namen: Ethernet- bzw. Netz-Adresse, physische Adresse...
 - Beispiel: 00:00:A7:10:86:96
- Hersteller von Netzkarten kaufen einen Adressblock und vergeben eindeutige Adressen (zweite 3-Byte-Gruppe)
- Geeignete Netzkartentreiber können ursprüngliche MAC-Adresse überschreiben
 - ermöglicht strukturierte lokale Adressen
- Es gibt einige spezielle (“funktionale”) Adressen
 - z.B. Gruppenadresse für alle Bridges; für den aktiven Ringmonitor etc.
- Netzkarte des Empfängers akzeptiert folgende Pakete (d.h. leitet sie an die höhere Schicht weiter):
 - Zieladresse = eigene Adresse
 - Zieladresse = Broadcast-Muster (alle Bits der Zieladresse = 1)
 - Zieladresse = eine Multicastadresse, auf die der Netzadapter des Empfängers programmiert wurde, und Multicast-Bit = 1 (Multicast = Teilmenge aller Stationen, z.B. alle Dateiserver)
 - alle Pakete, falls der Empfänger im “promiscuous-Modus” ist (Netzadapter lässt sich in diesen Modus versetzen, z.B. um Fehler in einem LAN zu analysieren oder um Bridges und Gateways realisieren zu können --> beliebter Modus für Hacker...)

Transceiver

- Sende- und Empfangslogik
- Kollisionsentdeckung

Transmitter / Receiver



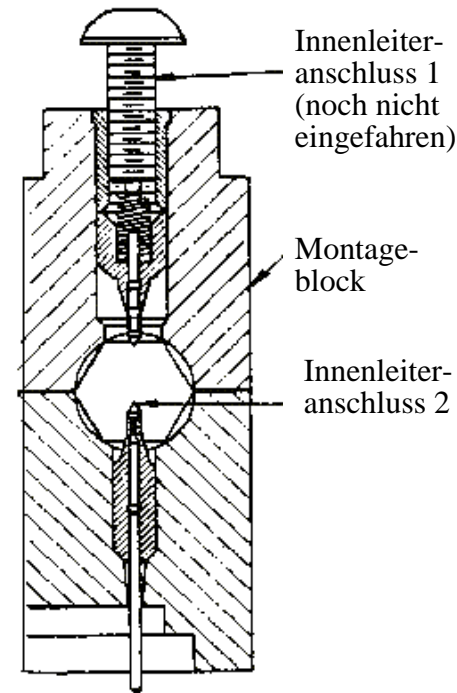
- Transceiver-Kabel aus (mind.) 9 einzelnen Leitungen
 - 2 × transmit; 2 × receive; 2 × collision presence; 2 × power; shield
 - 15-Pin-Stecker (“DIX”)

Frühe Transceiver

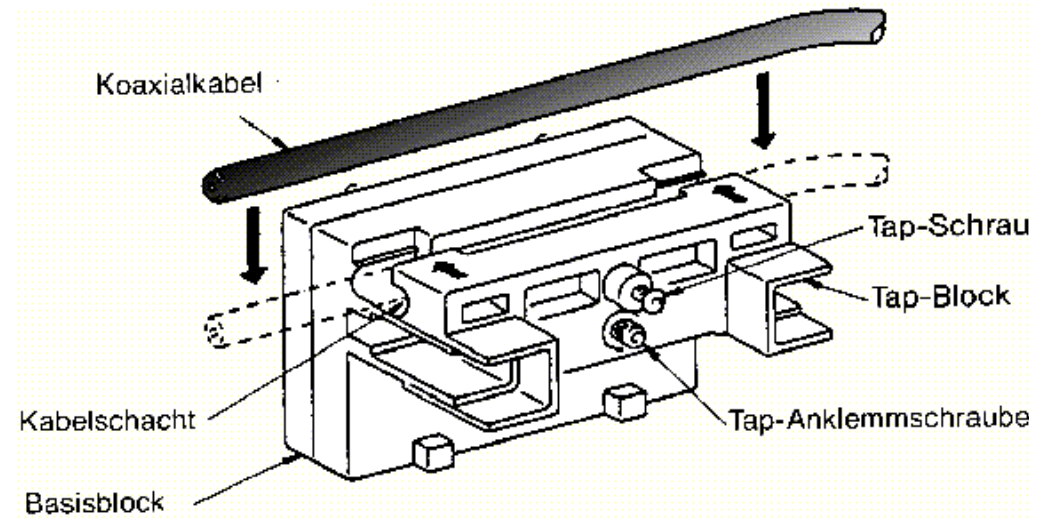
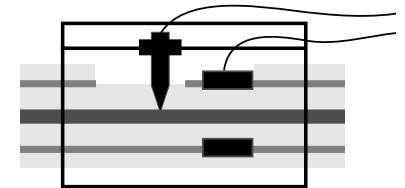


Ethernet-Anschluss mit TAP-Vampirklammer

Terminal Access Point

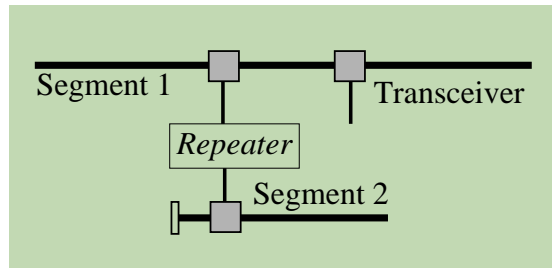


- spitze Sonde wird an den Innenleiter gepresst
- Klemme als Kontakt zum Aussenleiter



Repeater

- Bidirektionaler Signalverstärker auf dem Physical Layer
- Zweck: Koppeln bzw. Segmentieren von Ethernets

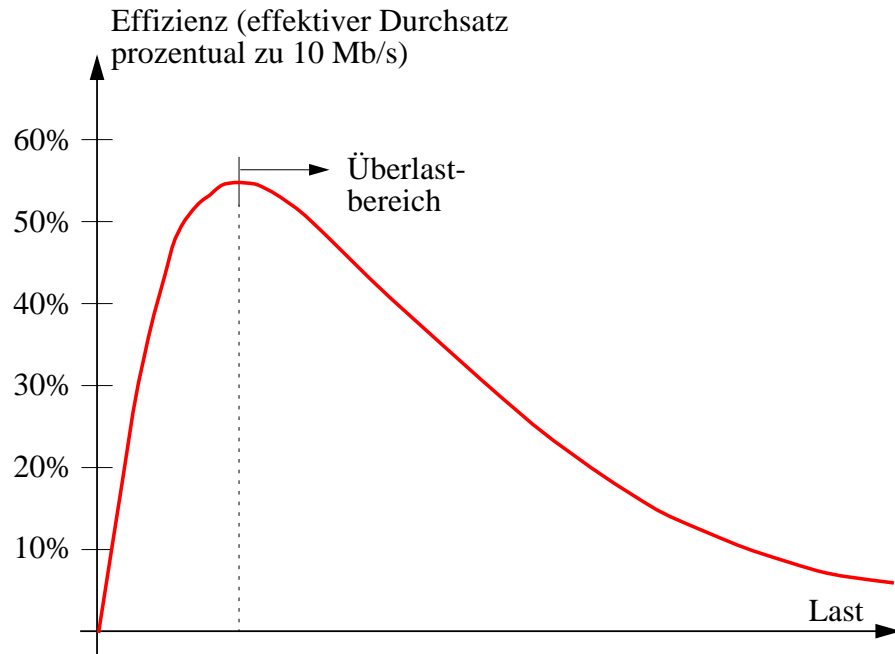


- Aufgabe u.a.:
 - taktgerechte Signalregenerierung
 - Kollisionserkennung
 - Erzeugung von Jam-Signalen
 - Abtrennen fehlerhafter Segmente (z.B. bei Dauerkollisionen wegen fehlerhaftem Abschlusswiderstand)
- Verzögerung sollte kleiner als ca. 8 Bitzeiten sein
 - entspricht wenigen hundert Meter Kabel
- Beachte beim Koppeln von Ethernet-Segmenten:
 - es dürfen keine Schleifen erzeugt werden (Topologie also: Baum)
 - Multisegment-Ethernet bildet eine einzige "Kollisionsdomäne"

Ethernet: Stabilität, Durchsatz

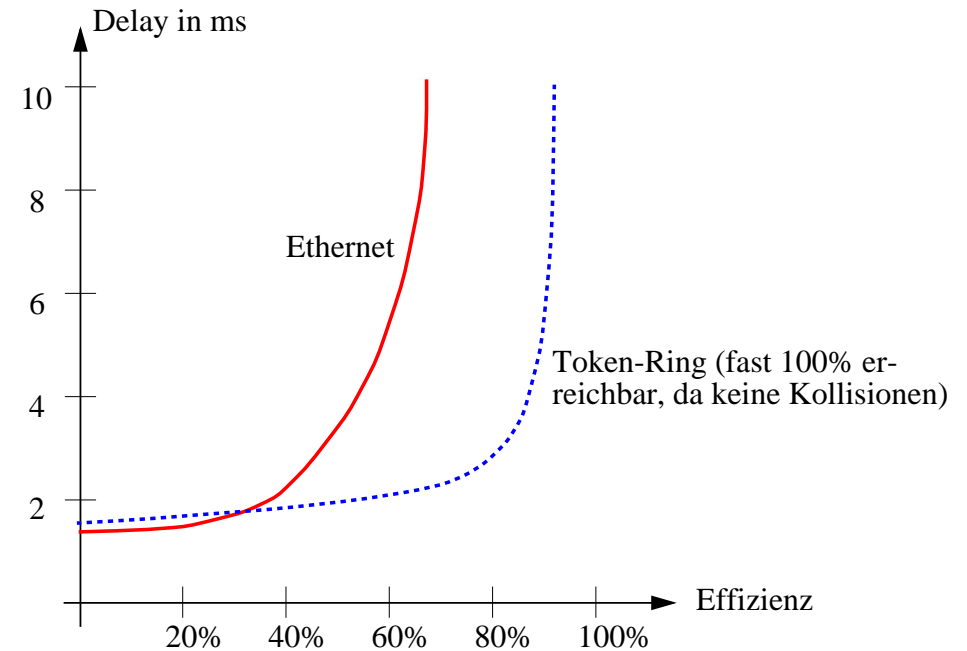
- Ein Paket kann "zufällig" beliebig lange verzögert werden
 - Abbruch nach einer bestimmten Anzahl von Versuchen (--> "Fehler")
 - Nicht für ("harte") Realzeitanwendungen geeignet
 - Zugangsverfahren nicht fair oder deterministisch
- Bei hoher Eingangslast wird das System instabil
 - Hohe Kollisionszahl } --> sinkender Gesamtdurchsatz
 - Viele Wiederholungen } (Vgl. Staus auf Autobahn bei Überlast)
 - Höchstens 30-50% Auslastung (sonst überlastet)
- Durchsatz von 10 Mb/s gilt nur bei schwacher Gesamtauslastung (wenig Sender, wenig Pakete auf dem Netz)
 - Typischerweise können die 10 Mb/s nur etwa zur Hälfte effektiv genutzt werden (Rest: Kollisionen, Wartezeiten, Retransmissionen; geringers Problem, wenn man das Netz für sich alleine hat)
 - Theoretische Bandbreite wird ausserdem beschränkt durch:
 - Flusssteuerung höherer Protokolle
 - Overhead durch zu kleine Pakete (Paketkopf!)
 - Overhead der Kommunikationssoftware in den Stationen
 - Typische Ethernetsegmente: nicht mehr als 10 - 20 Stationen

Durchsatz und Überlastverhalten



- Typisches Verhalten, wenn mehrere Stationen unabhängig voneinander Last (= zu sendende Pakete) erzeugen
 - lässt sich analytisch, simulativ und experimentell ermitteln
- Der Überlastbereich ist typischerweise schon erreicht, wenn es zu 5 - 10% Paketwiederholungen kommt
 - bei mehr als 5% Kollisionen (aller gesendeten Pakete) sollten Massnahmen getroffen werden (z.B. Aufteilung des Netzes)

Delay bei Hochlast

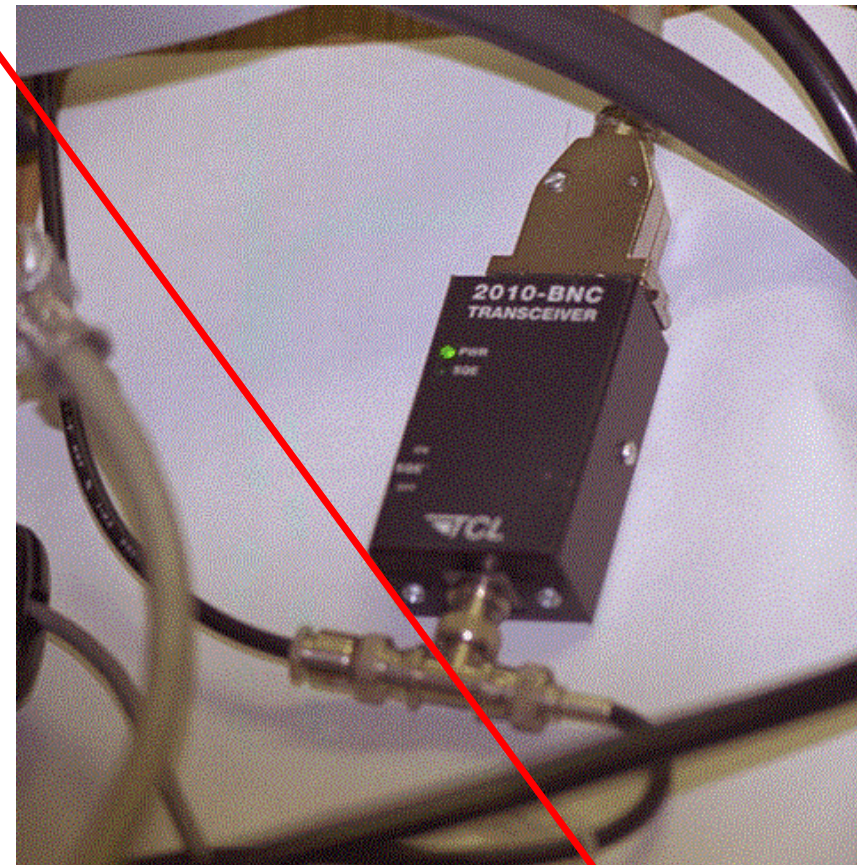
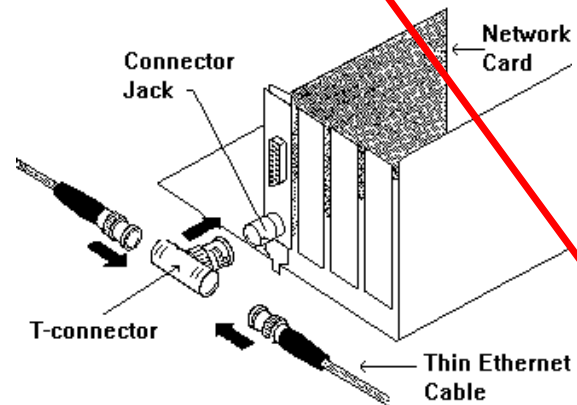


- Bei schwacher Last (< ca. 40% Auslastung) hat Ethernet gegenüber Token-Ring etwas geringere Verzögerung
 - beim Token-Ring muss man auch im "Leerlauf" auf das Token warten
- Oberhalb von 40%-50% sollte ein Ethernet nicht betrieben werden; dann steigt die Verzögerungszeit rasant

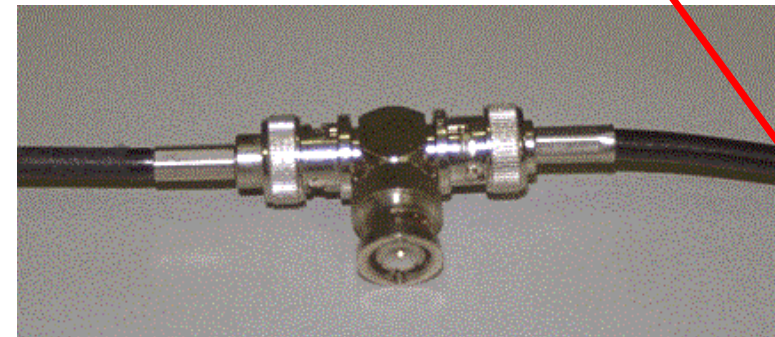
Thin-Wire-Ethernet

Auch als "Cheapernet" oder "10 Base 2" bezeichnet

- 4.8 mm Koaxkabel statt 10.2 mm
 - einfacher zu verlegen und preiswerter
- Segmentlänge max. 185 m statt 500 m
- Transceiver i.a. direkt auf der Netzkarte
- Anschluss über BNC-T-Stecker direkt am Rechner
 - "daisy-chain"
- Max. 30 Stationen pro Segment



Transceiver mit AUI-Schnittstelle für Thin-Ethernet

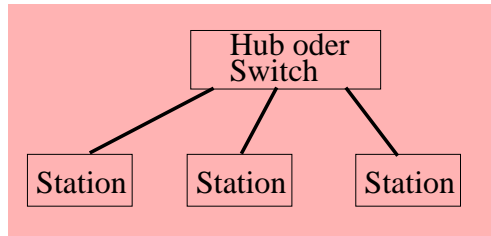


Direktanschluss über BNC-T-Stecker bei Thin-Ethernet

10 Base T - Ethernet

- Sternförmig mit twisted pair

- auch Punkt-zu-Punkt-Verbindungen zwischen 2 Stationen möglich



- 4-fach UTP, max 100 m
- RJ-45-Konnektor
- Transceiver i.a. direkt auf der Netzadapter-Karte

- Hub ist i.w. ein Multiport-Transceiver ("Konzentrator")

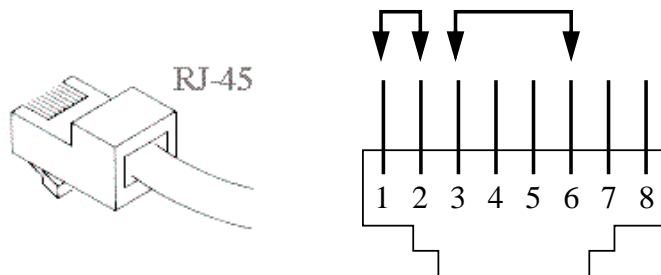
- generiert Kollisionssignal, wenn mehrere Stationen gleichzeitig senden
- statt Multiport-Transceiver oft auch ein "Switch" (--> später)

- Kabelstörungen betreffen i.a. nur eine Station

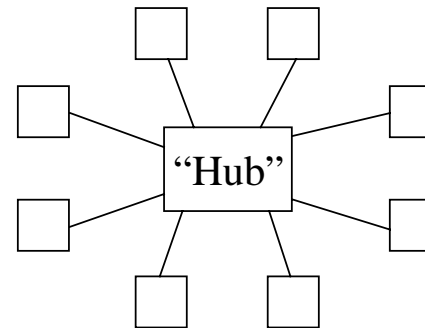
- Hub entfernt im Fehlerfall betroffene Station logisch
- ggf. auch periodischer Test und automatische Reaktivierung

- RJ-45-Anschluss

- bei 10 Base T-Ethernet sind 4 von 8 Anschlüssen benutzt



Hub ("Sternkoppler")



- "Hub" = Zentrum, "Dreh- und Angelpunkt"... (vgl. auch zentrale Umsteige-flughäfen von Airlines)
- sternförmige Verkabelung (innerhalb einer Etage oder Arbeitsgruppe)

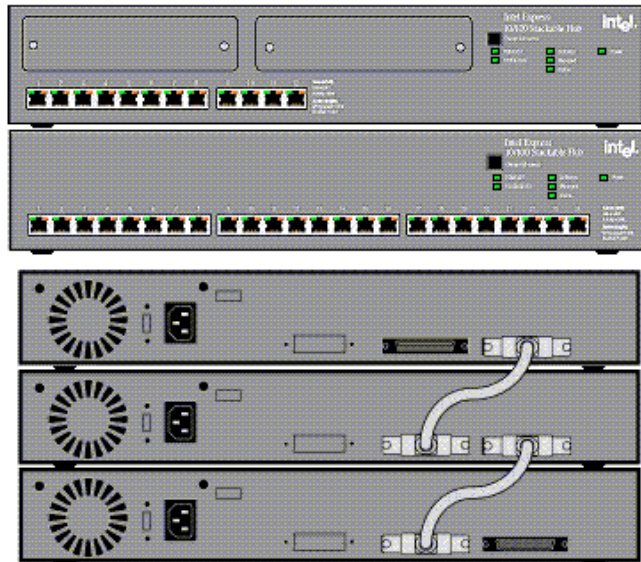
- Diverse Ausprägungen: von einfachen Multiport-Repeater ("klassischer Hub") bis zu Switching-Funktionalitäten

- einfache Hubs bilden bei Ethernet i.a. eine einzige "Kollisionsdomäne" (d.h.: physischer Stern, aber logisch ein einziger Bus!)
- oft: Hub kennt die MAC-Adresse des einzigen an einem Port angeschlossenen Gerätes und leitet nur solche Pakete dort hin
- enthält oft zusätzliche Funktionen für Netz-Management

- Konfigurierbar über verschiedene Einschubkarten, z.B.

- einige 10 Mb/s-Ethernet-Ports oder 100 Mb/s-Ethernet-Ports
- "uplink-Anschluss" für ein schnelles Backbone-Netz

Hub-Ansichten



The Hub's Manual

Not long ago, the computer network at the offices of American Scientist developed a frustrating problem - turtle-paced file transfers between machines and refusals to print documents, all punctuated by unpredictable periods of normal function. A long diagnosis isolated the problem to a piece of hardware called a **hub**, essentially a connecting device where computers, printers, scanners and so on meet to communicate. While attempting to verify the device's "illness," one of our editors turned to the **hub's manual, which offered one rather useless bit of diagnostic information: The activity light should be flashing randomly.** So I watched the pinhead-size green light flash on and off-blinking in a seemingly erratic pattern. It looked random enough, but then I wondered: *What is random?*

...

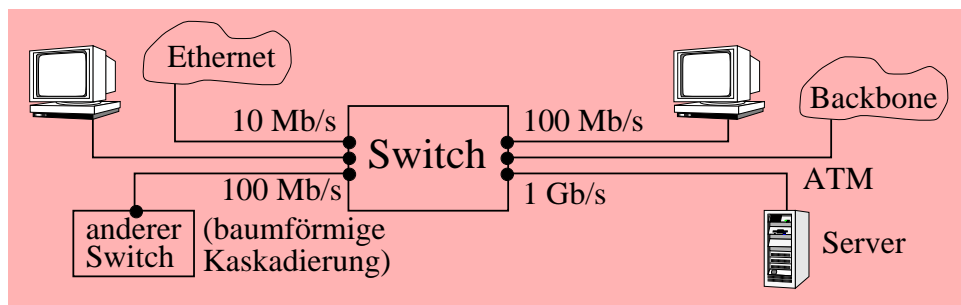
<http://www.amsci.org/amsci/issues/Sciobs97/Sciobs97-05Random.html>



Linkes Bild: Router und Hub im linken Rack; Patchfeld für Lichtwellenleiter im rechten Rack; *rechtes Bild:* Router und Hub mit Twisted-Pair-Kabeln

LAN-Switching

- Dedizierte Verbindungen zwischen je zwei Ports
 - ähnlich wie das Herstellen einer Telefonverbindung
 - idealerweise mehrere Verbindungen gleichzeitig ohne Beeinflussung
 - Einschränkung von Fehlern auf den Verursacher; ggf. logische Trennung (Deaktivierung) des Ports
 - leistungsfähige Backplane notwendig
 - sternförmige Verkabelung (i.a. twisted pair)



- Ethernet-Switch: simuliert auf jeder Leitung ein eigenes Ethernet (Koppelung wie Bridges auf Ebene 2)
 - Endgeräte unverändert (wie im klassischen Ethernet) nutzbar
 - jeweils dediziert volle Bandbreite pro Leitung
 - i.a. voll Duplex; keine "unnötigen" Kollisionen } höhere Gesamtleistung
- Switch i.a. ausbaufähig (verschiedene Einschubkarten)
 - andere Protokolle
 - höhere Geschwindigkeiten für einzelne Leitungen
 - mehr Funktionalität (z.B. Management; Statistiken; Sicherheit)
- Kritische Komponente --> Fehlertoleranz notwendig
 - redundante Netzteile, Lüfter...; Austausch im laufenden Betrieb

LAN-Switching (2)

- Weiterleiten der Pakete auf zwei Arten:
 - a) Store and forward:
 - Paket wird erst ganz empfangen
 - dadurch Prüfsumme ermittelbar; fehlerhafte Pakete nicht weiterleiten
 - b) Cut through (bzw. "on the fly bridging"):
 - Weitervermittlung beginnt sofort nach Erkennen der Zieladresse (spart bei 10 Mb/s zwischen 14 μ s und 1.2 ms, je nach Paketlänge)
 - volle Filterfunktionen (z.B. Trigger auf bestimmte Paketinhalte bzgl. bestimmter Bitmuster höherer Protokolle) nicht möglich
 - Kompromiss: Zumindest abwarten, bis Mindestpaketlänge erreicht (damit werden kurze Reste abgebrochener Pakete herausgefiltert)

- Flusststeuerung

- bei einem belegten Zielport muss ein Paket ggf. (kurz!) gepuffert werden, oder es wird dem Sender ein Kollisionssignal geschickt
- Puffer pro Eingangsport, pro Zielport, oder pro Pfad (= zugeordnete Eingangs- / Zielports)
- Staus können sich bei belegtem Zielport ergeben, oder z.B. wenn 100 Mb/s am Eingang auf 10 Mb/s am Ausgang umgesetzt werden muss
- "nicht blockierend": eine Verbindung von a nach b blockiert nicht das Weiterleiten von Paketen von c nach d (aufwendiger --> teurer)

- Stapelbare ("stackable") Switche

- Koppelung des Busses ("Backplane") i.w. gleichartiger Geräte
- effizienter als Kaskadierung (baumförmige Hintereinanderschaltung)

- Adressauflösung (Zuordnung MAC-Adresse zu Port)

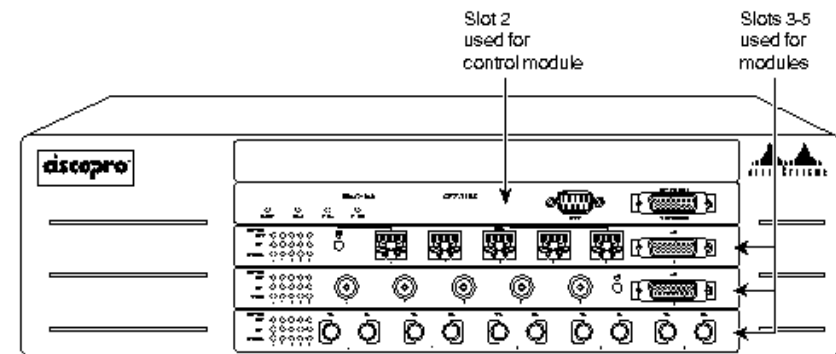
- ggf. statisch konfiguriert bei einem einzelnen Gerät am Port (Sicherheit!)
- oder dynamisches Lernverfahren wie bei Brücken

Switch-Ansichten

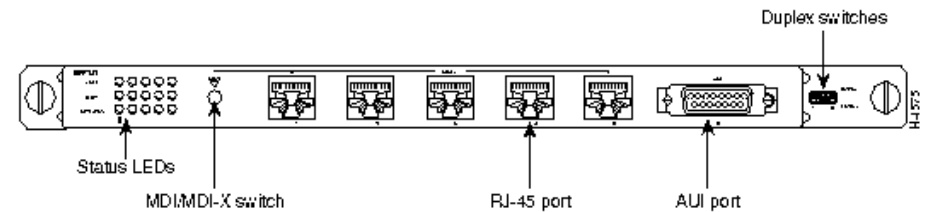


Ein Beispiel: Cisco EtherSwitch CPW2115

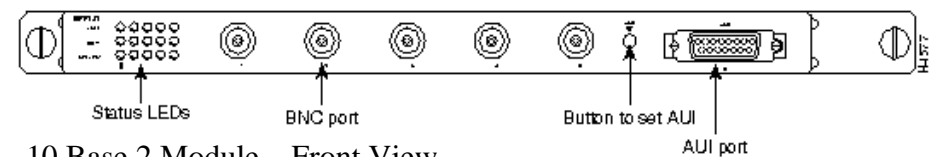
The switch increases network throughput by interconnecting devices via dedicated 10-Mbps ports over multiple media types including 10 Base T, 10 Base 2, and 10 Base FL. Designed specifically for networks built on mixed media, it provides MultiMAC connections between hubs, switches, routers, and servers through its 15-port capacity. It has a modular chassis that accommodates any combination of three 5-port, hot-swappable expansion cards.



CiscoPro EtherSwitch CPW2115---Front View



10 Base T Module---Front View



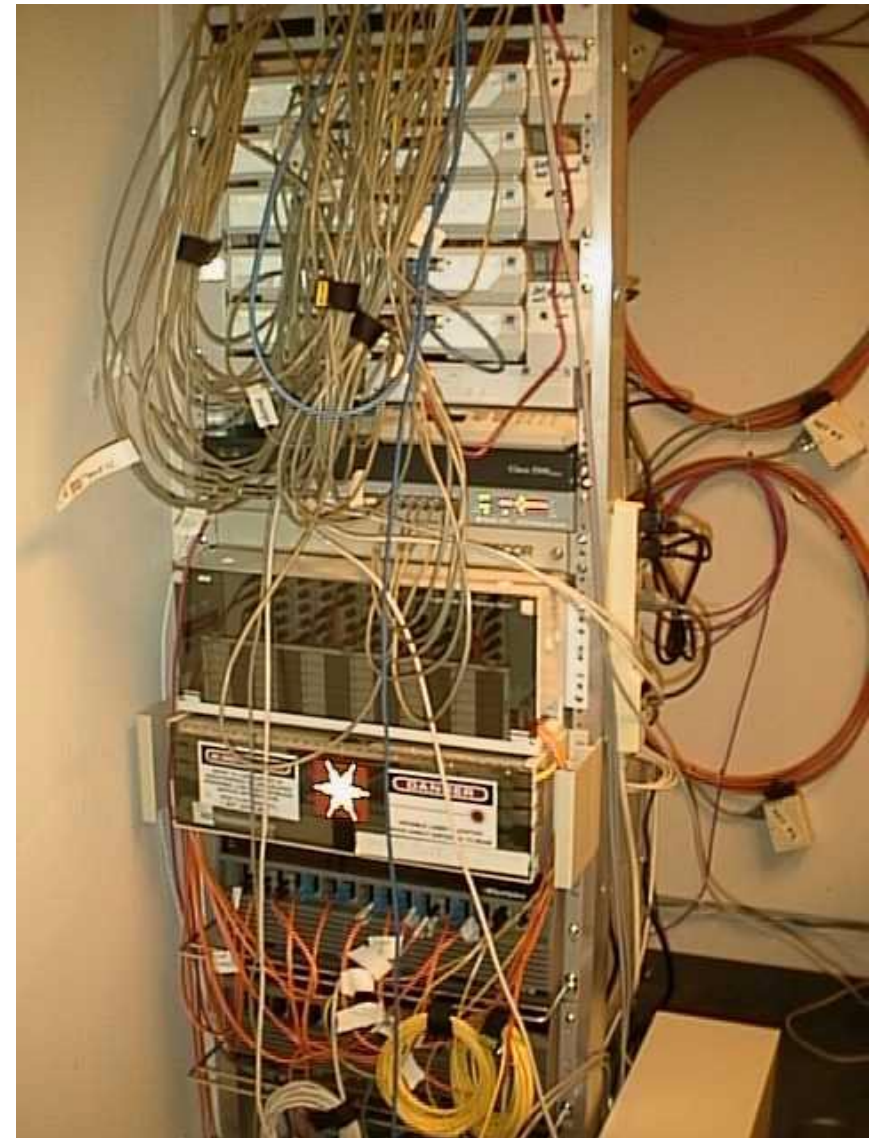
10 Base 2 Module---Front View

Daten zum EtherSwitch CPW2115

- Auszug aus der Produkt-Spezifikation

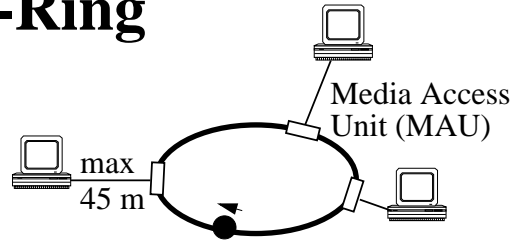
Filter rate	223,200 packets per second
Forward rate	23,200 packets per second
Routing delay	40 microseconds from first bit at input port to first bit at output port
Addresses per port	1,700 active addresses per port
	6,000 addresses per system = 4-MB memory
	10,000 addresses per system = 8-MB memory
Buffers	256 full-size packets per port
Number of ports	15: any combination of three 5-port modules for 10 Base T, 10 Base 2, or 10 Base FL. On 10 Base T and 10 Base 2 modules, 1 AUI port may be used in place of 1 other port.

Switching Equipment



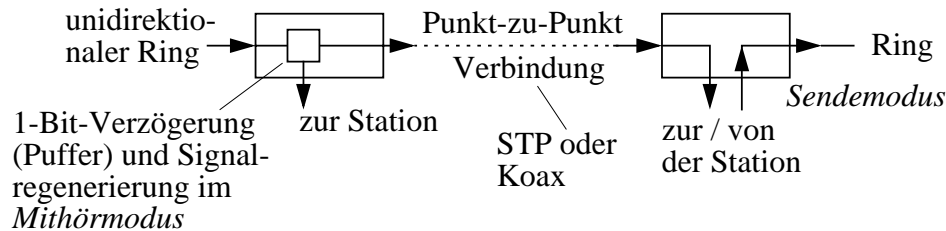
Token-Ring

- Protokoll genormt: IEEE 802.5



- Routing nicht nötig
 - Broadcast möglich
- } Nachricht umrundet Ring ganz

- Sendemodus (max. eine Station) / Mithörmodus



- Freitoken umrundet Ring

- Sendewillige Station:

- wartet auf Freitoken
- ersetzt Freitoken durch Belegttoken mit angehängter Nachricht
- fügt Empfangsadresse an den Anfang der Nachricht

i.a. ≤ 5000 Byte

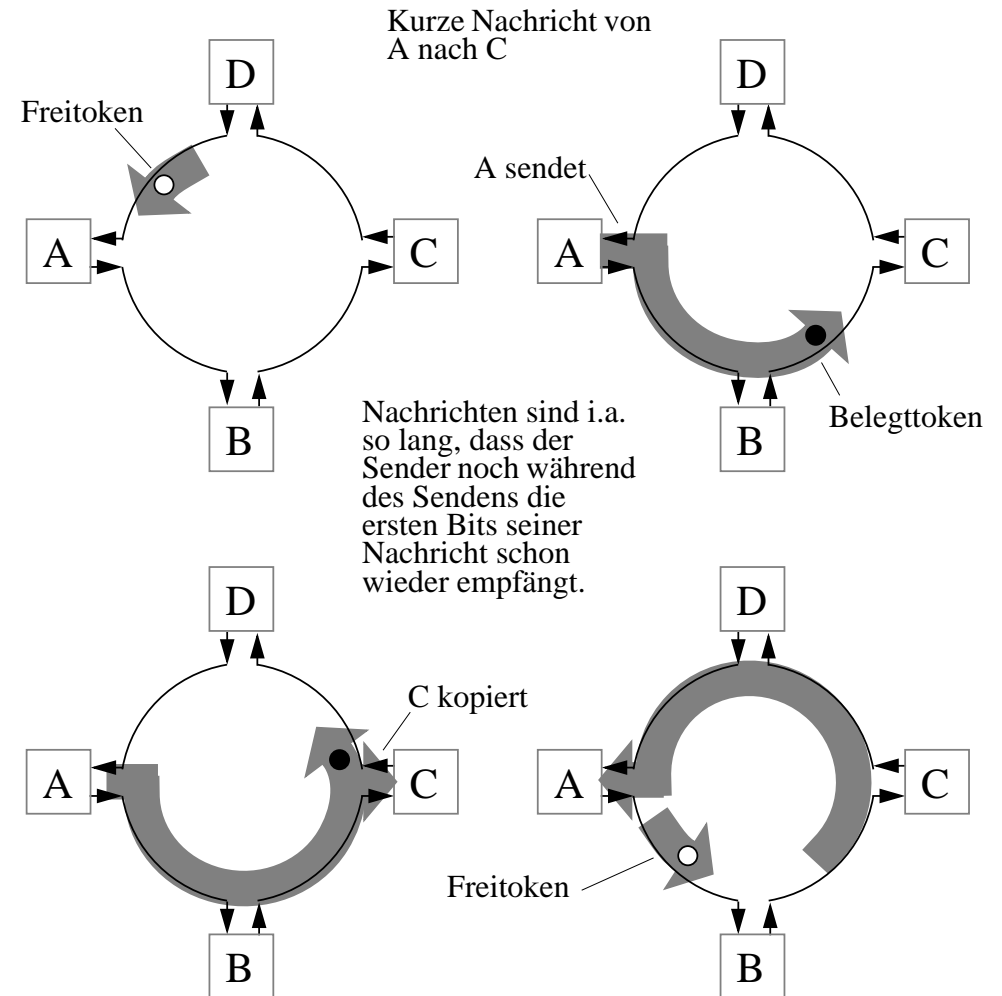
- Nachricht:

- wird vom Empfänger gelesen
- umrundet Ring weiter, wird dann vom Sender wieder absorbiert

- Token:

- Belegttoken nach Umrundung vom Sender in Freitoken verwandelt --> zirkulierendes Senderecht
- Timeout für maximale "token holding time" (typw. 10 ms)

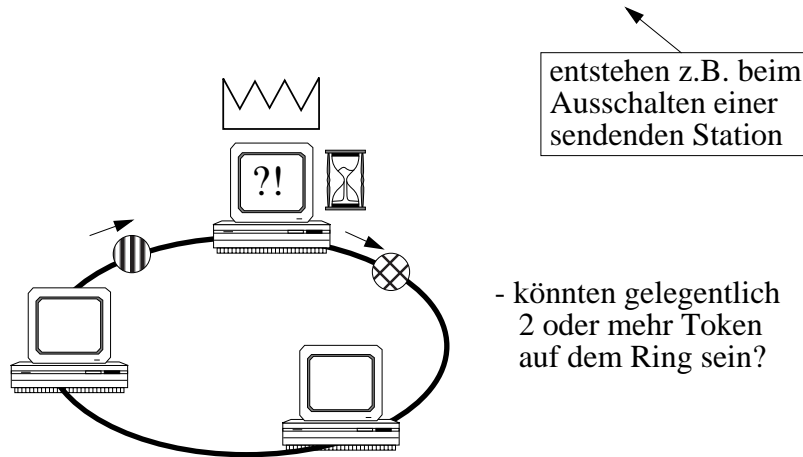
Token-Ring: Datenübertragung



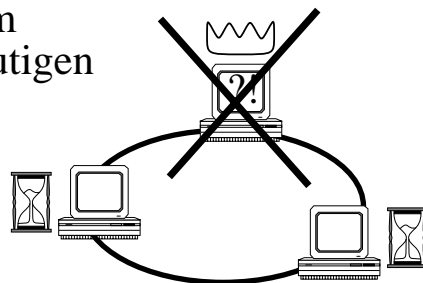
- Token-Ring wurde Anfang der 80er-Jahre vom IBM-Forschungslabor in Rüschlikon entwickelt
- Hat heute gegenüber den schnellen Ethernet-Varianten stark an Bedeutung verloren

Der Monitor

- Eine Station übernimmt die Rolle des *Monitors*
- > Erkennung und Behebung von Fehlersituationen
 - timeout, um verlorene Token zu erkennen (--> regenerieren)
 - stempelt jedes Token, um mehrfaches Zirkulieren zu erkennen
 - absorbiert alte Token und fehlerhafte Nachrichtenreste



- *Monitorausfall* wird von anderen Stationen durch einen “grossen” Timeout erkannt (lange weder Token noch “Active Monitor Present”-Nachricht gesehen)
- > Election-Protokoll, um *dezentral* einen eindeutigen Monitor zu wählen



Election-Protokoll wird auch benötigt, damit die Stationen nach “erstmaliger” Konfiguration einen Monitor auswählen

Der Monitor-Election-Algorithmus

- Hier nur das *Prinzip*: In der Praxis komplizierter durch diverse Timeouts, Fehlerbedingungen etc.
- Voraussetzung: Identitäten aller Stationen sind verschieden
- Jede Station mit Identität p hat eine lokale Variable M
- M ist initial 0; am Ende enthält M die Identität des Monitors

Adresse der Station

Bewerber mit der höchsten Identität

Jede Station p führt folgenden Algorithmus aus:

Bei einem Timeout:

```
M := p;
send <M> to next
```

Bei Eintreffen einer Nachricht $\langle j \rangle$:

```
if j > M then /* Wahl verloren...*/
{ M := j;
  send <M> to next
}
if j = p then "Wahl gewonnen!"
```