## 3. Link Layer

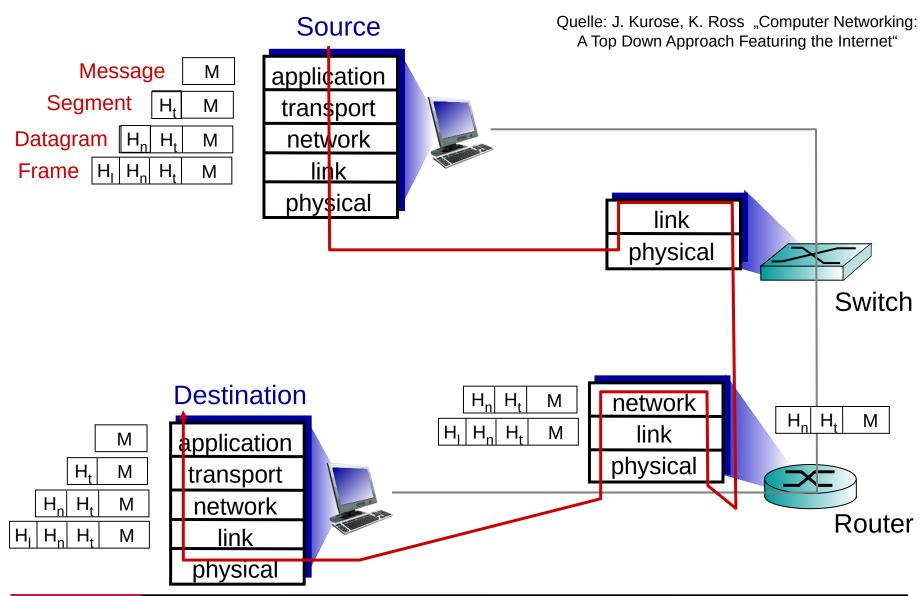
- 3.1 Übersicht
- 3.2 LANs
  - 3.2.1 Adressierung
  - 3.2.2 Ethernet
  - 3.2.3 Hubs und Switches

Some slides are adapted from Computer Networking: A Top Down Approach Featuring the Internet, 3rd - 6rd edition. Jim Kurose, Keith Ross, Addison-Wesley. All material copyright 1996-2012, J.F Kurose and K.W. Ross, All Rights Reserved





### Datenfluss im Schichtenmodell

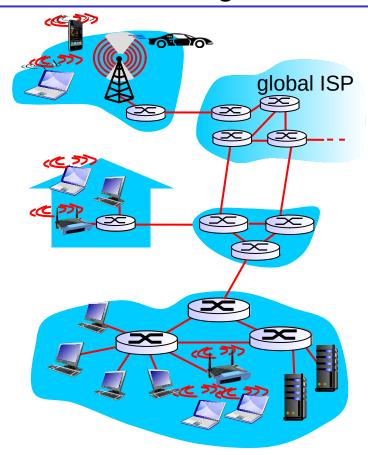




Die "Data"-Link Layer ist verantwortlich für die Übertragung von Datagrammen von einem Knoten zu einem **physisch** benachbarten Knoten über eine Verbindung

## Terminologie:

- Knoten (Nodes): Hosts & Router
- Links(Verbindungen):
   Kommunikationskanäle die benachbarte Knoten auf einem Kommunikationspfad verbinden
  - Kabelgebundene / kabellose Verbindungen
  - LANs (können Switches beinhalten)
- Frame (Rahmen): Layer-2-Paket
  - enthält Datagramm (L3)







## Link Layer Services

### Framing

 Kapselt ein Datagramm in einen Frame, fügt Header (H) und Trailer (T) hinzu

Frame H Datagramm T

#### Link Access

- Medium access control (MAC) wird benötigt wenn es sich um einen geteilten (shared) Kanal handelt
- "MAC"-Adressen werden in Frame-Headern genutzt um die Quelle (src) und das Ziel (dst) zu identifizieren
  - Unterscheiden sich von IP-Adressen!
- Half-duplex und full-duplex
  - Bei Halbduplex können Knoten auf beiden Seiten der Verbindung senden, jedoch nicht gleichzeitig (z.B. Sprache im Raum)





## **Link Layer Services**

Verlässliche Zustellung zwischen benchbarten Knoten

### Fehlererkennung

- Fehler werden durch Signaldämpfung und Störungen (Rauschen) verursacht
- Empfänger erkennt Existenz von Fehlern und signalisiert dem Sender, dass erneute Übertragung nötig ist, oder verwirft den Frame

#### Fehlerkorrektur

- Empfänger identifiziert und korrigiert Bitfehler ohne Neuübertragung
- Selten auf Verbindungen mit geringer Fehlerrate (fiber, twisted pair)
- Wird auf Drahtlosverbindungen mit hoher Fehlerrate genutzt

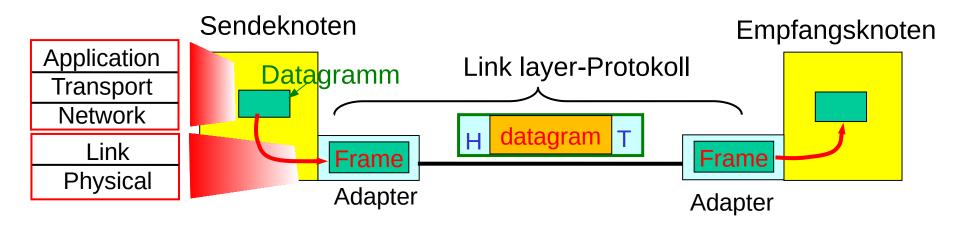
#### Flow control

- Mechanismus der verhindert, dass der empfangende Knoten Daten von einem benchbarten, sendenden Knoten schneller erhält als er verarbeiten kann
- Ist für Layer 2 durch ISO-OSI Protokolle definiert





## Link Layer Services – Implementation Issues



- Link layer ist in Ethernet "Adapterkarte" (NIC) implementiert
  - Adapter ist semi-autonomous und implementiert link & physical layers
- Senderseite:
  - Kapselt das Datagramm in Frame
  - Fügt Header (H) und Trailer (T) bits hinzu um: src/dst-Adressen zu identifizieren Error checking bits, etc. hinzuzufügen
- Empfängerseite
  - Fehlererkennung oder Korrektur
  - extrahiert Datagramm,
     übergibt an Empfangsknoten
  - flow control (ISO-OSI)





## Link Layer and Interconnection Networks

- Point-to-point links
  - PPP für Einwählverbindungen (weitgehend obsolet)
  - point-to-point link zwischen Ethernet-Switch und Host
- Broadcast links (geteiltes Kabel oder Medium)
  - Old-fashioned Ethernet
  - 802.11 wireless LAN
  - upstream HFC (=Fernsehkabel, DOCSIS)



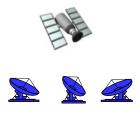
humans at a cocktail party (shared air, acoustical)



shared wire (e.g., cabled Ethernet)



shared RF (e.g., 802.11 WiFi)



shared RF (satellite)





# 3. Link Layer

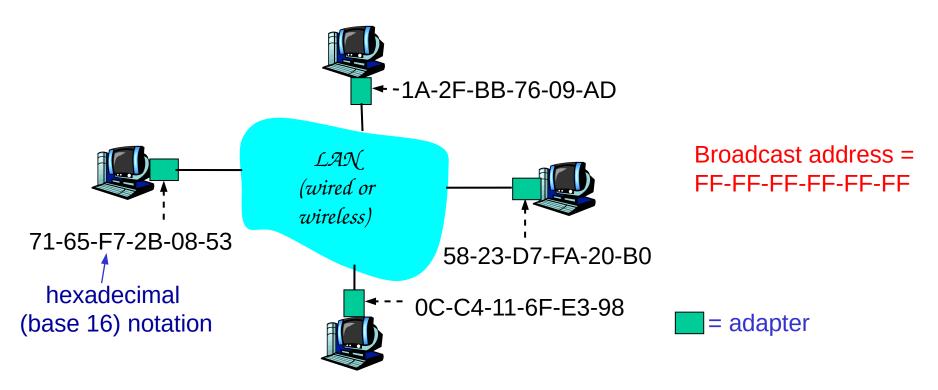
- 3.1 Übersich
- 3.2 LANs
  - 3.2.1 Addressierung
  - 3.2.2 Ethernet
  - 3.2.3 Hubs und Switches





### MAC Adressen

- MAC (oder LAN oder physical oder Ethernet)-Addresse
  - Genutzt um Frame von einem Interface zu einem anderen physisch verbundenen Interface zu übertragen (selbes L2-Netzwerk)
  - 48 Bit MAC-Adresse (6 Byte) "burned" in Adapter-ROM
  - Jeder Adapter im LAN hat eine eigene einzigartige LAN-Address

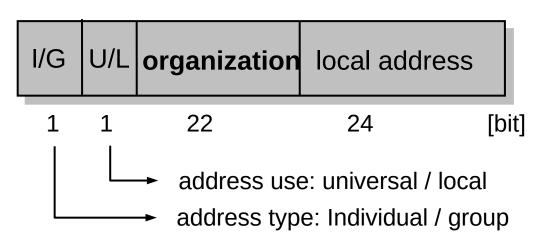






### MAC Adressen - 2

- Globale Adressen
  - Gesamtlänge 48 Bit (EUI-48); neu: EUI-64 (24+40 bit)
- MAC-Adressverteilung wird von IEEE administriert
- Hersteller kauft ein Stück MAC-Adressraum (um Einzigartigkeit der Adressen zu garantieren)
- MAC flat address → portability
  - LAN-Karte kann von einem LAN zu einem anderen bewegt werden während die selbe MAC-Adresse benutzt wird







# 3. Link Layer

- 3.1 Übersicht
- 3.2 LANs
  - 3.2.1 Adressierung
  - 3.2.2 Ethernet
  - 3.2.3 Hubs und Switches



## Unzuverlässiger, Verbindungsloser Dienst

### Verbindungslos:

 Kein Verbindungsaufbau oder handshaking zwischen sendendem und empfangendem Adapter

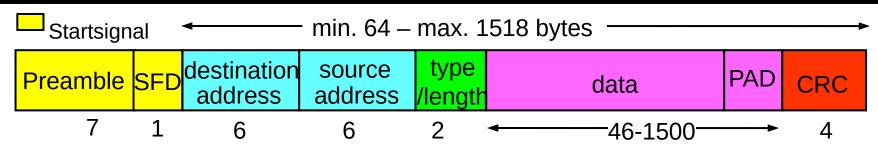
### Unzuverlässig:

- Empfangender Adapter sendet keine ACKs oder NACKs zum sendenden Adapter
- Keine Neuübertragung wird ausgelöst wenn der empfangene Rahmen Bitfehler aufweist (falsche CRC, Cyclic Redundancy Check)
- Neuübertragung wird aber im Falle von Kollisionen ausgelöst
- Stream von Datagrammen der an die Network Layer weitergegeben wird kann Lücken aufweisen
- MAC-Protokoll
  - Ethernet nutzt CSMA/CD mit Exponential-Backoff-Retransmission Algorithmus





### Struktur von Ethernet Frames - IEEE 802.3



#### Preamble and SFD

• 7 Byte Preamble (10101010) wird genutzt um die Taktraten zu synchronisieren, gefolgt von einem Byte (10101011) Start of Frame Delimiter (SFD).

#### Adressen:

• Wenn Adapter einen Frame mit passender Zieladresse oder mit Broadcastadresse erhält, wird der Frame verarbeitet, ansonsten wird der Frame verworfen

#### Type/length

- Zeigt das Protokoll der höheren Schicht an (e.g. IP, Novell IPX, AppleTalk)
- length gibt die Anzahl an Bytes im Data-Feld an

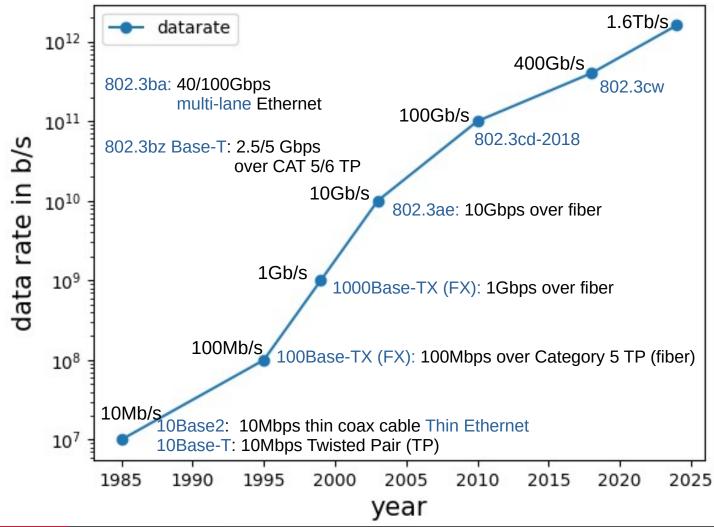
#### ■ PAD:

- Wird hinzugefügt wenn weniger als 46 Byte an Daten vorhanden sind
- Minimale Frame-Länge von 512 Bit/64 Bytes wird für Collision detection benötigt
- CRC (Cyclic Redundancy Check):
  - Wird am Empfänger geprüft, Frame wird verworfen wenn Fehler erkannt wird





# **Ethernet Standardisierung**

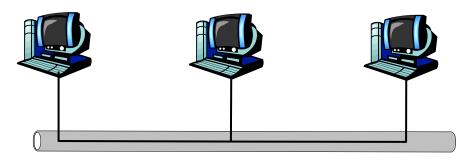




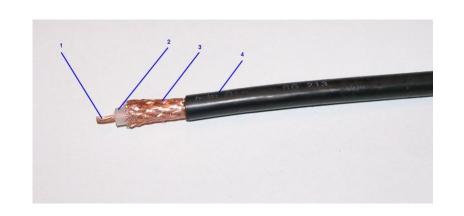
## **Ethernet Bus Topologie**

#### 10Base5 und 10Base2

- 10 Mbps Datenrate
- Bus-Topologie weit verbreitet in den 90ern
- Nutzt Koaxialkabel zur Geräteverbindung
- Um die 2500mMaximaldistanz zwischenEndgeräten
- MAC-Protokoll:
   Carrier Sense Multiple Access with Collision Detection (CSMA/CD)



**Ethernet** 



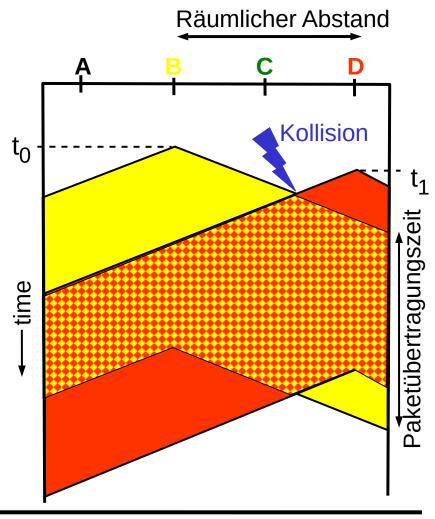


# Carrier Sense Multiple Access (CSMA)

### **CSMA – Prinzip**

- "Hören" vor dem Senden:
- Wenn Kanal als belegt erkannt wird, Übertragung verschieben
- Wenn Kanal als frei erkannt: Übertragung des gesamten Rahmens
- Kollisionen können trotzdem passieren:
  - Dies liegt an der propagation delay, da zwei Knoten ihre Übertragungen unter Umständen nicht hören können
  - Gesamte Übertragungszeit ist verschwendet

### Räumliches Layout der Knoten

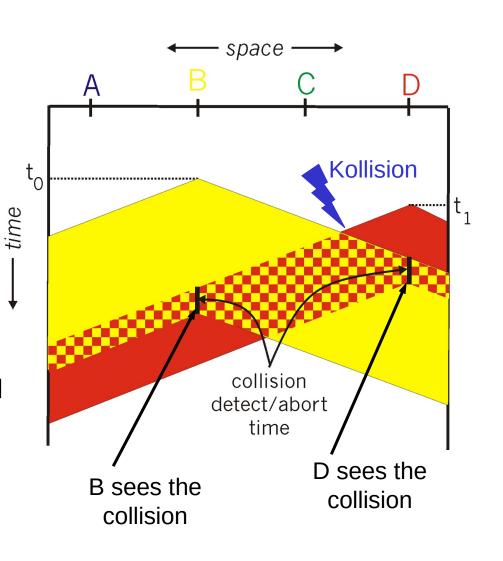






### **CSMA/CD Collision Detection**

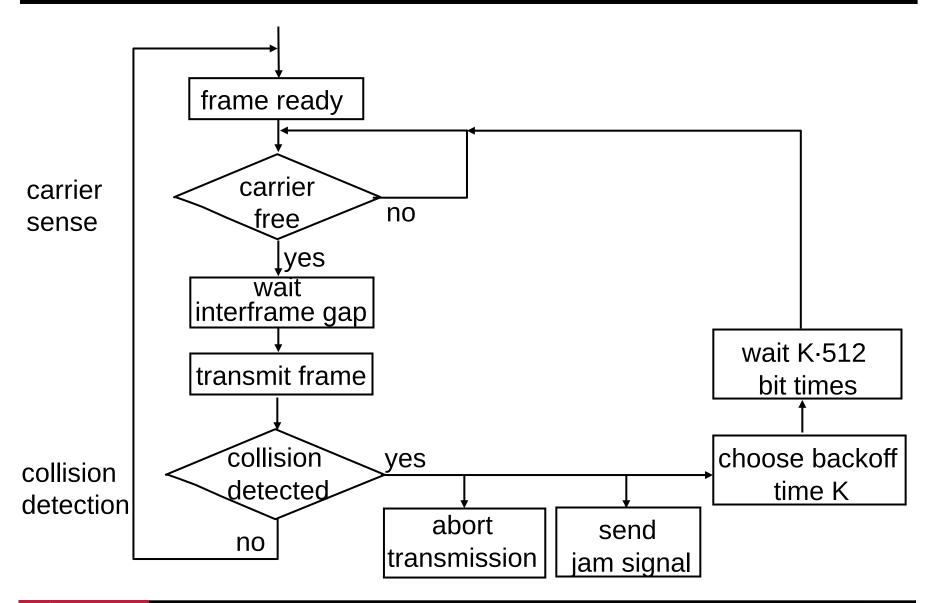
- Kollisionserkennung (CD)
  - Wenn Kollision innerhalb kurzer Zeit erkannt wird, werden die kollidierenden Übertragungen abgebrochen
  - Hieraus folgt eine reduzierte "Kanalverschwendung"
- Implementierung von CD
  - Einfach in kabelgebundenen LANs: messen der Signalstärke, vergleichen des gesendeten und empfangenen Signals
  - Schwierig in drahtlosen LANs: Sender wird während des Sendens abgeschaltet







# Ethernet CSMA/CD Algorithmus







## Ethernet CSMA/CD Algorithmus

- Keine Synchronisation, z.B. Time Slots
  - 1. Adapter empfängt Datagramm vom Netzwerk-Layer und erzeugt ein Frame, z.B. gespeichert im Puffer
- Carrier Sense
  - 2. Wenn der Adapter sieht, dass der Kanal "idle" ist, beginnt er mit der Übertragung des Frames nach dem Interframe Gap (IFG)
    - Sieht er, dass der Kanal belegt ist, wartet er bis der Kanal wieder "idle" ist.
    - Der Interframe Gap beträgt 96 Bit Zeiten (9.6µs für 10Mbit/s LAN) um vorangegangene Zustände zu bearbeiten (Wechsel zwischen Senden und Empfangen)
    - Der Interframe Gap stellt ebenfalls Fairness zwischen den Stationen sicher.
  - 3. Wenn der Adapter ein gesamtes Frame übertragen hat, ohne eine andere Übertragung zu detektieren, ist dieses erfolgreich abgeschlossen.





## Ethernet CSMA/CD Algorithmus

#### Collision Detection

- 4. Detektiert der Adapter eine andere Übertragung während er selber sendet, bricht er diese ab uns sendet ein Jam Signal.
  - Das Jam Signal stellt sicher, dass alle anderen Teilnehmer von der Kollision erfahren; 48 Bit groß
  - Die minimale Framelänge (512 Bit) ist für Collision Detection notwendig
    - Stellt sicher, dass die Signaldauer größer ist als die Round Trip Time
- Retransmission Strategie
  - 5. Nach Abbruch der Übertragung wechselt der Adapter zum Exponential Back-Off:
    - nach der m-ten Kollision wählt der Adapter ein zufälliges K aus dem Intervall {0, 1, 2, ..., 2<sup>m</sup>-1}
    - Bemerke: m ist auf m ≤ 10 begrenzt
  - 6. Adapter wartet K · 512 Bit Zeiten und kehrt zu Schritt 2 zurück.





## **Exponential Back-Off**

#### Ziel:

- Passt die Anzahl der Übertragungswiederholungen an die geschätzte aktuelle Netzwerkauslastung an.
  - Starke Auslastung: zufällige Wartezeit länger
- Erste Kollision: wähle K aus {0, 1}; Delay ist K · 512 Bitübetragungszeiten
- Nach der zweiten Kollision: wähle K aus {0, 1, 2, 3}, ...
- Nach der zehnten Kollision: wähle K aus {0, 1, 2, 3, ..., 1023}

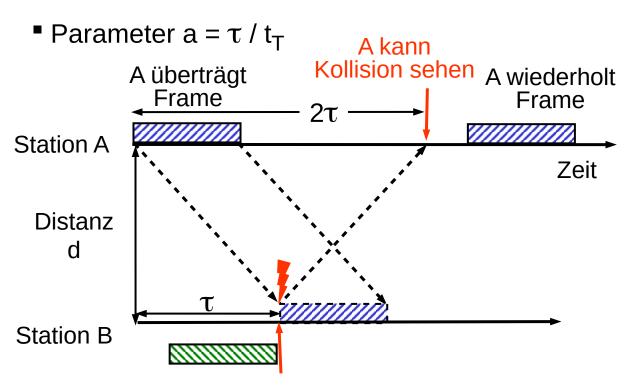
#### Nachteile:

- Für 10Mbps Ethernet ist die Bitzeit 0.1µs
  - → für K = 1023 beträgt die Wartezeit ca. 50ms



## Worst-case collision detection in Bus Topologien

- v<sub>B</sub> = Kanalübertragungsrate [bit/s]
- $\mathbf{T}$  = maximale propagation delay zwischen 2 Knoten im LAN
- $\mathbf{t}_{T} = \mathbf{L} / \mathbf{v}_{B}$  Zeit um maximalgroßen Frame der Länge L [bits] zu übertragen



B erkennt Kanal als frei kurz bevor der Frame ankommt und fängt an zu senden → Kollision

Slot time

Worst-case collision detection time

$$T_{SL} = 2\tau$$

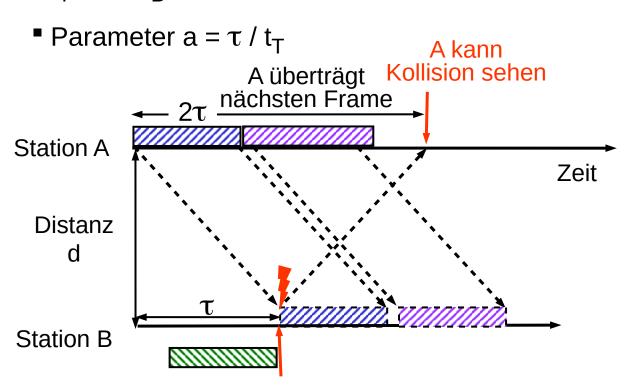
Frage: In welchem Fall schlägt dieser Ansatz fehl?





## Worst-case collision detection in Bus Topologien

- v<sub>B</sub> = Kanalübertragungsrate [bit/s]
- $\tau$  = maximale propagation delay zwischen 2 Knoten im LAN
- t<sub>T</sub> = L / v<sub>B</sub> Zeit um maximalgroßen Frame der Länge L [bits] zu übertragen



A kann nicht feststellen ob der erste oder der zweite Rahmen kollidiert ist.

Problemanalyse: die im Beispiel genutzte Framelänge ist zu kurz für CSMA/CD

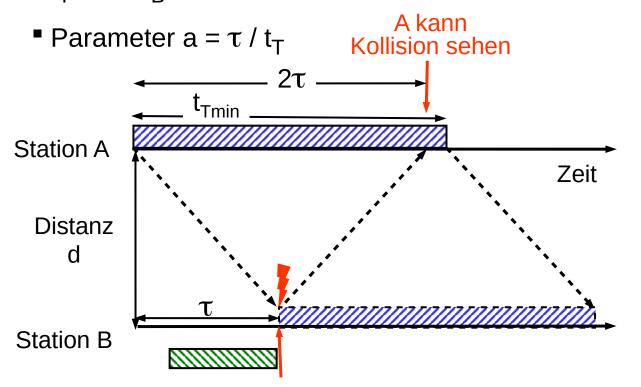
B erkennt Kanal als frei kurz bevor der Frame ankommt und fängt an zu senden → Kollision





## Worst-case collision detection in Bus Topologien

- v<sub>B</sub> = Kanalübertragungsrate [bit/s]
- $\tau$  = maximale propagation delay zwischen 2 Knoten im LAN
- t<sub>T</sub> = L / v<sub>B</sub> Zeit um maximalgroßen Frame der Länge L [bits] zu übertragen



A kann nicht feststellen ob der erste oder der zweite Rahmen kollidiert ist.

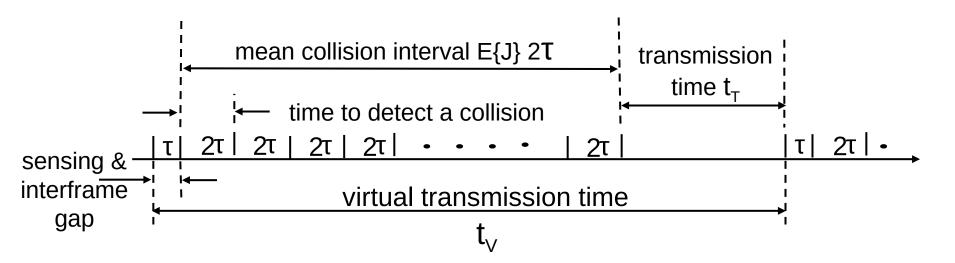
Problemanalyse: die im Beispiel genutzte Framelänge ist zu kurz für CSMA/CD

B erkennt Kanal als frei kurz bevor der Frame ankommt und fängt an zu senden → Kollision





## Throughput in Bus Topologien



- Mittleres Kollisionsintervall E{J}
  - E{J} ist die durchschnittliche Anzahl an Übertragungswiederholungen, vorausgesetzt, dass Kollisionen aufgetreten sind.
- Virtuelle Übertragungsdauer
  - Dauer um ein Frame erfolgreich zu übertragen, wenn Kollsionen aufgetreten sind:

$$t_{V} = t_{T} + \tau + 2 \tau \cdot E\{J\} = t_{T} \cdot \left(1 + \frac{\tau}{t_{T}} + \frac{2\tau}{t_{T}} \cdot E\{J\}\right) = t_{T} \left(1 + a + 2a \cdot E\{J\}\right)$$





## Erfolgswahrscheinlichkeit

#### **Annahmen:**

- Im Netzwerk sind N >> 1 aktive Stationen
- Jede Station hat immer eine Dateneinheit zum übertragen
- Sei dir Wahrscheinlichkeit, dass eine beliebige Station in einem 2τ Intervall übertragen möchte gleich p.
  - Die Wahrscheinlichkeit, dass k von N Stationen übertragt ist dann

$$P_{k} = \frac{N}{k} p^{k} (1-p)^{N-k}$$

• Die Wahrscheinlichkeit, dass genau eine Station überträgt (erfolgreiche Übertragung) ist

$$\hat{p} = P_1 = Np(1-p)^{N-1}$$

Der Wert p = 1/N maximiert die Erfolgswahrscheinlichkeit

$$\hat{p}_{max} = \left(1 - \frac{1}{N}\right)^{N-1} \xrightarrow{N \to \infty} e^{-1}$$



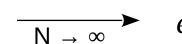
## Länge des Kollisionsintervalls J

- Keine Kollisionen: j = 0
  - Nur eine Station sendet im Intervall 2au mit der Wahrscheinlichkeit  $\hat{p}_{max}$
- Eine Kollision ist aufgetreten (j > 0); der Exponential Back-Off Algorithmus wird wie folgt approximiert:
  - Das Kollisionsintervall ist eine Einheit (2 $\tau$ ) lang wenn während dieses Intervalls genau eine Station ihre Übertragung wiederholt (Wahrscheinlichkeit  $P(j=1)=\hat{p}_{max}$ )
  - Ist das Intervall 2 Einheiten lang:  $P\left(j=2\right) = \hat{p}_{max}\left(1 \hat{p}_{max}\right)$
  - Ist das Intervall j Einheiten lang:  $P(J=j) = \hat{p}_{max} (1 \hat{p}_{max})^{j-1}$

(Geometrische Verteilung)

Durchschnittliches Kollisionsintervall:

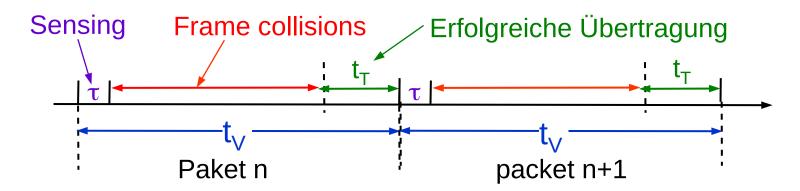
$$E\{J\} = \sum_{j=0}^{\infty} j \cdot P\{J = j\} = \sum_{j=1}^{\infty} j \cdot \hat{p}_{max} (1 - \hat{p}_{max})^{j-1} = \frac{1}{\hat{p}_{max}}$$







## Maximum Throughput and Utilization



■ Tatsächliche Paketübertragungszeit t<sub>v</sub>

$$\simeq$$
 Sensing + collision time + transmission time  $t_V = \tau + 2\tau \cdot e + t_T = t_T (1 + a(1 + 2e))$  (a =  $\tau / t_T$ )

Maximale Paketdurchsatzrate

$$\lambda_{\text{max}} = \frac{1}{t_V}$$
 [Pakete pro Sekunde (pps)]

Maximale Auslastung

$$\rho_{\text{max}} = \frac{t_{\text{T}}}{t_{\text{V}}} = \frac{1}{1 + a(1 + 2e)} \approx \frac{1}{1 + 6.44a}$$

Geht gegen 1 wenn  $\tau$  gegen 0 geht





## Beispiel

### Bus Topologie 10 Mbit/s Ethernet

- N = 40 Stationen
- d = 2500m;  $v_A = 2.3 \times 10^8 \text{ m/s} \rightarrow \tau = d / v_A \approx 11 \text{ ms}$
- L = 2200 bit;  $v_R = 10 \text{ Mbit/s} \rightarrow t_T = L / v_B = 220 \text{ ms}$
- $a = \tau / t_{\tau} = 0.05$
- Maximale Auslastung:  $\rho_{\text{max}} \approx \frac{1}{1 + 6.44a} = 0.756$
- effektive Bitrate (für alle N Stationen):  $v_{max} = v_B \rho_{max} = 7.56$  Mbit/s
- maximale mittlere Bitrate pro Station:  $v_N = v_{max} / N = 189$  kbit/s
- Upgrade des Bus auf 100 Mbit/s Ethernet
  - L = 2200 bit;  $v_B = 100 \text{ Mbit/s} \rightarrow t_T = L / v_B = 22 \text{ ms}$
  - $a = \tau / t_T = 0.5$ ,  $\rho_{max} = 0.237$
  - $v_{max} = v_B \rho_{max} = 23.7 \text{ Mbit/s}$
  - $v_N = 592 \text{ kBit/s}$





# 3. Link Layer

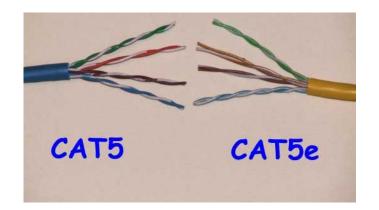
- 3.1 Übersicht
- 3.2 LANs
  - 3.2.1 Addressierung
  - 3.2.2 Ethernet
  - 3.2.3 Hubs und Switches

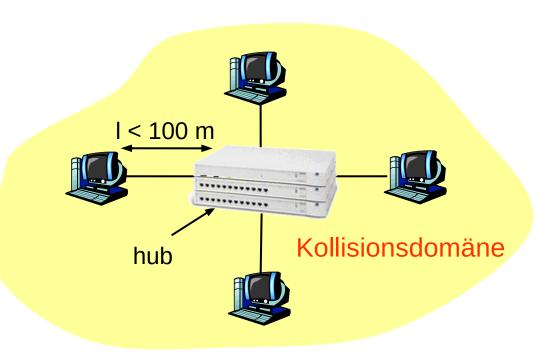




#### 10BaseT und 100BaseT

- Rate von 10/100 Mbps, später als "Fast Ethernet" bezeichnet
- T steht für Twisted Pair
- Knoten verbunden mit einem Hub:
  - "Sterntopologie"; 100 m maximale Entfernung zwischen Knoten und Hub
  - → 5m pro Seite für Anschlussleitung, 90m Kabel dazwischen
- Alle Hosts sind in der selben Kollisionsdomäne (denn Hub=Layer1, terminiert Layer1 also nicht)









### Hubs sind de-facto physical-layer Repeater

- Bits kommen auf einer Verbindung an und werden auf allen anderen Verbindungen wiederholt
  - Mit der selben Rate
  - Kein Frame buffering (kein Zwischenspeichern)
- Kein CSMA/CD am Hub
  - Kollisionen treten auf wenn zwei Frames zur selben Zeit ankommen

• Adapter erkennen nur Kollisionen

ankommender Frame

ausgehender Frame

Layer 1 Repeater Layer 1





## Beispiel Sterntopologie

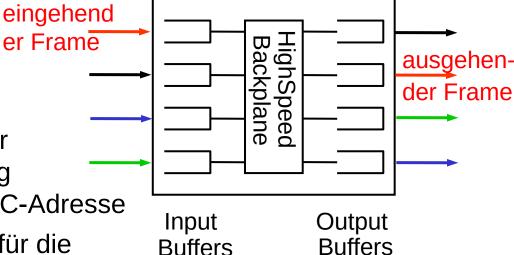
- Sterntopologie 100 Mbit/s Ethernet
  - N = 40 Stationen
  - d = 2 \* I = 200m;  $v_A = 2.3 \times 10^8 \text{ m/s} \rightarrow \tau = d / v_A \approx 1 \text{ ms}$
  - L = 2200 bit;  $v_B = 100$  Mbit/s  $\rightarrow$   $t_T = L / v_B = 22$  ms
  - $a = \tau / t_T = 0.0455$
  - maximale Auslastung:  $\rho_{max} = 0.773$
  - effektive Bitrate (für alle N Stationen):  $v_{max} = v_B \rho_{max} = 77.3 \text{ Mbit/s}$
  - maximale mittlere Bitrate pro Station:  $v_N = v_{max} / N = 1.93$  Mbit/s
- Welchen Preis hat der höhere Druchsatz?
  - Zusätzliches zentrales Gerät: der Hub
  - Teurere Verkabelung und deutlich kürzere Distanz (200m)

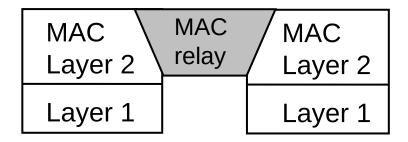




### Ethernet Switches – Verbessern der Performance

- Ein Ethernet Switch ist ein Paket-Switch auf der Link Layer (L2)
  - "Stores and forwards"
     Ethernet Frames
  - Buffering von Frames verhindert Kollisionen
  - Auswerten der Frame Header und selektive Weiterleitung basierend auf der Ziel-MAC-Adresse
  - Nutzt das Ethernet-Protokoll für die Verbindungen der Interfaces
  - Unterstützt heterogene Verbindungen
    - Verschiedene Medien auf Layer 1
    - Kombinationen von geteilten/ dedizierten
       10/100/1000 Mbps-Interfaces

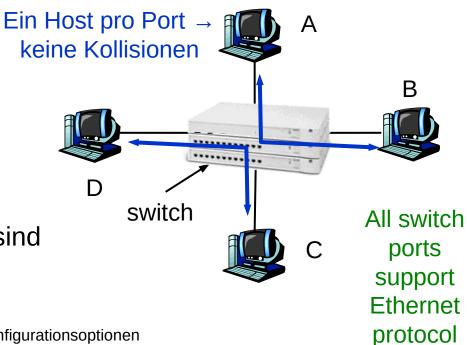






### Stern-Netzwerk mit Ethernet Switch

- Jedes Interface ist isoliert und bildet seine eigene Kollisionsdomäne
- Wenn Hosts dedizierten Zugang zum Switch haben:
  - Vollduplexübertragung mit Leitungsgeschwindigkeit zwischen zwei Hosts
  - Beispiel; switching: A-zu-B und C-zu-D simultan, keine Kollisionen
- Ein Switch ist transparent
  - Hosts wissen nichts von der Existenz des Switches
- Plug-and-play, self-learning
  - Switches benötigen keine Konfiguration\*
  - Switch lernt welche Hosts über welches Interface erreichbar sind



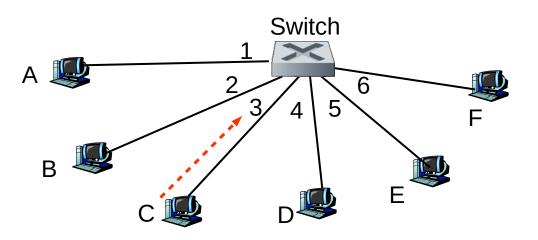
\*Manageable Switches können über eine Reihe an Konfigurationsoptionen verfügen, sind jedoch nicht Gegenstand der Vorlesung





## Self-Learning

- Ein Switch verfügt über eine Hardware Address Table (HAT oder "L2 Forwarding Table") mit folgenden Einträgen:
  - (MAC Address, Interface, Time Stamp)
  - Abgelaufene Einträge werden gelöscht (TTL kann bis zu 60 min sein)
- Switch erlernt die Einträge und speichert sie in der Address Table
- Beispiel: Switch empfängt Frame von C auf Interface 3
  - sieht C's Quelladresse in Frame
  - Fügt C's Quelladresse zur Address Table als "learned" Zieladresse, die über Interface 1 erreicht werden kann, hinzu



#### **Hardware Address Table**

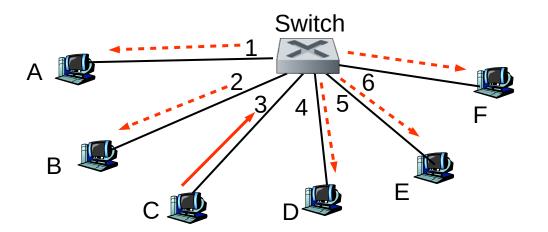
MAC-Addr	Interface	TTL
С	3	60





## Forwarding Beispiel

#### C sendet Frame zu D



MAC-add	Interface
CA	3 1
В	2
E	5

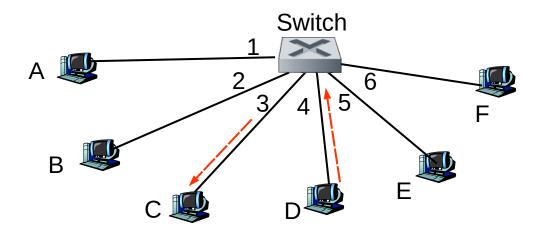
- Switch erhält Frame von C
  - Vermerkt in Address Table, dass C an Interface 3 verbunden ist
  - Weil D nicht in der Table ist, leitet der Switch den Frame zu den Interfaces 1, 2, 4, 5 und 6 mit Ausnahme des Empfangs-Ports (flooding)
- Frame wird von D empfangen





## Forwarding Beispiel

#### D antwortet C



MAC-add	Interface
С	3
Α	1
В	2
Е	5
ı-► D	4

- Switch erhält Frame von D
  - Vermerkt, dass D über Interface 4 erreichbar ist-
  - Weil C schon in der Table ist, leitet der Switch den Frame nur zu Interface 3
- Frame wird von C empfangen





## Self-Learning

Wenn Switch einen Frame empfängt:

Durchsuche Address Table nach Ziel-MAC-Adresse

if Eintrag für Ziel gefunden

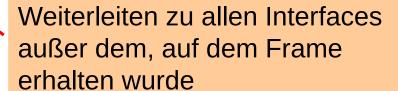
then{

if Ziel am selben Interface auf dem Frame erhalten wurdethen Frame verwerfen

else weiterleiten des Frames zum angegebenen Interface

}

else flood







### Verbinden von LANs über Switches

- Installation von Switches bricht große LANs in LAN-Segmente
  - Segmente werden separate Kollisionsdomänen
  - Heutzutage bestehen diese Segmente auch wieder aus Switches
- Switch filtert Pakete auf Basis der MAC-Adresse
  - Verkehr innerhalb des selben Segments wird normalerweise nicht in andere LAN-Segmente weitergeleitet (außer bei unbekannter MAC), e.g. HTTP/TCP-basierter end-to-end traffic
  - Traffic isolation zwischen verschiedenen Segmenten

