

Advanced computer networking

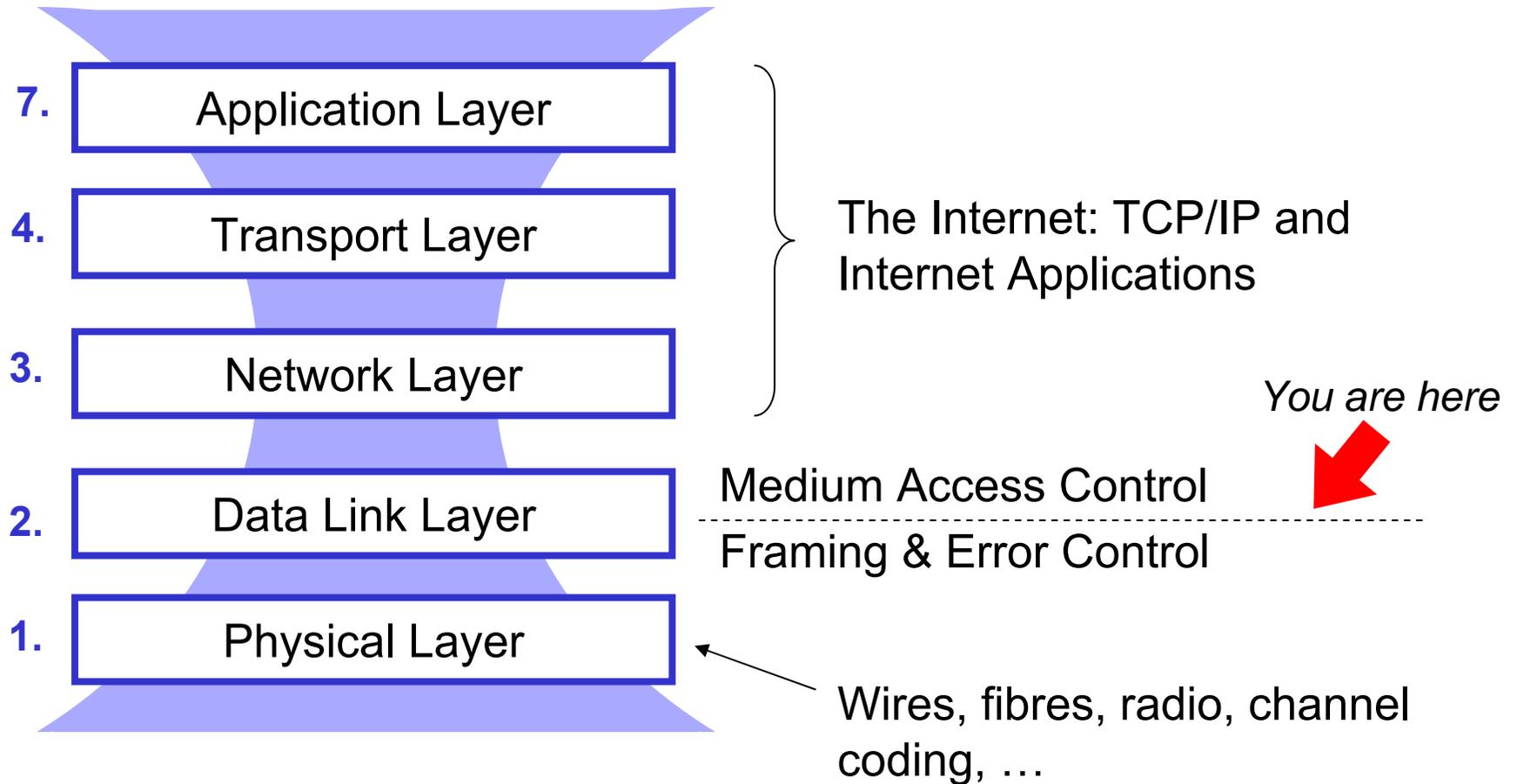
Internet Protocols

Thomas Fuhrmann



Network Architectures
Computer Science Department
Technical University Munich

Recap: The Data Link Layer

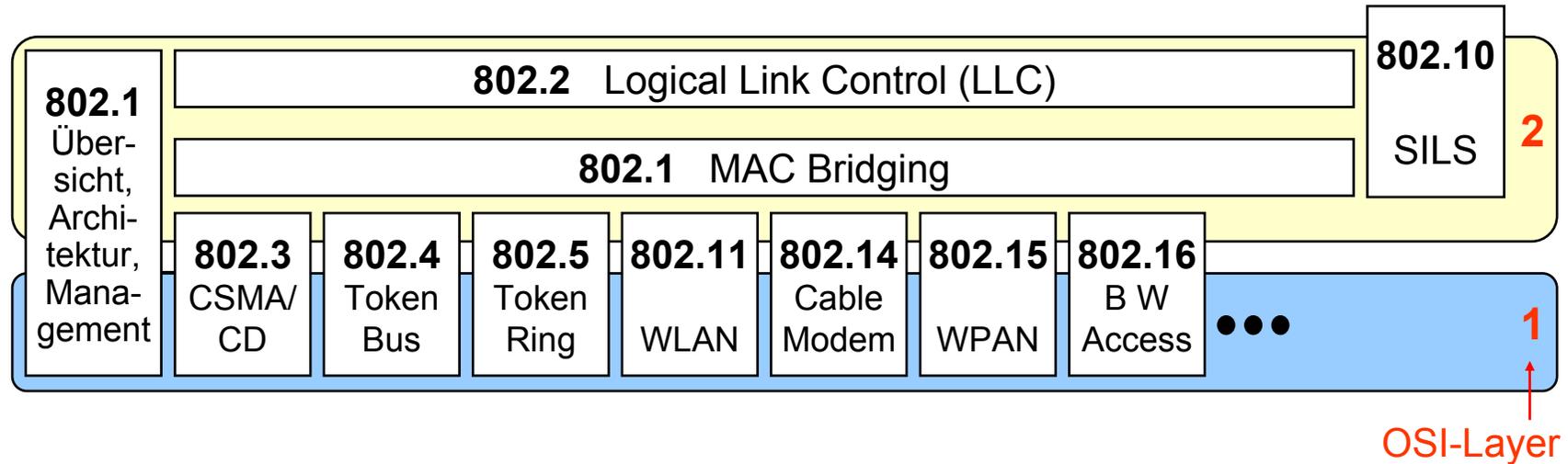


Medium Access Control Sub-Layer

- Many link layer technologies use a shared medium.
Examples: Ethernet, Wireless LAN, ...
- Multiple stations may want to send their frames almost at the same time, but the physical channel can only transmit one frame at a time.
 - How do I assign the channel to the stations?
 - Can I do this in a safe and fair manner?
- All stations will hear all the frames in the channel.
 - How do I address a particular station?

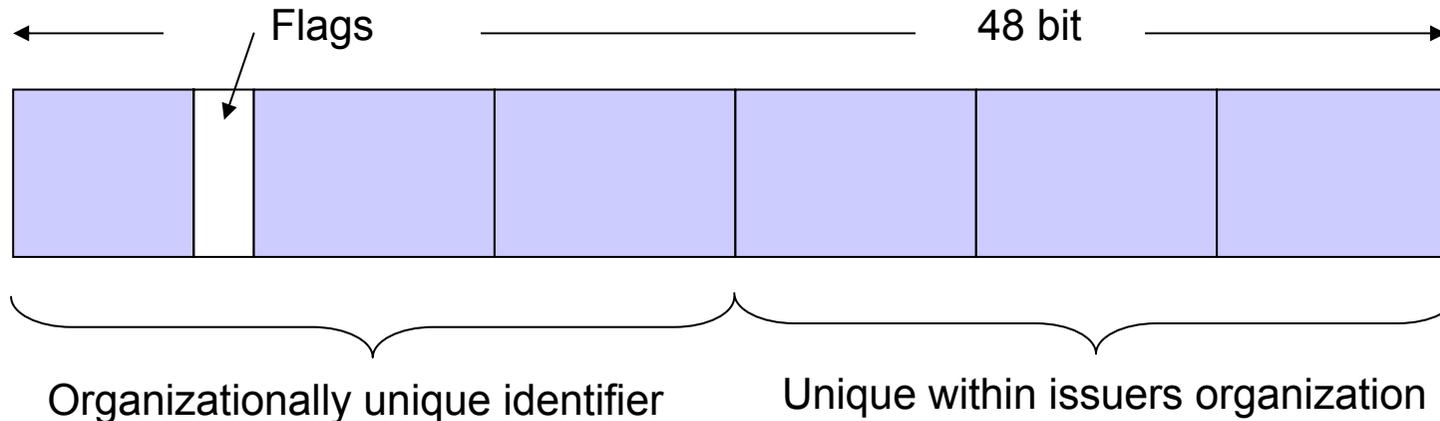


The IEEE 802 Protocol Family



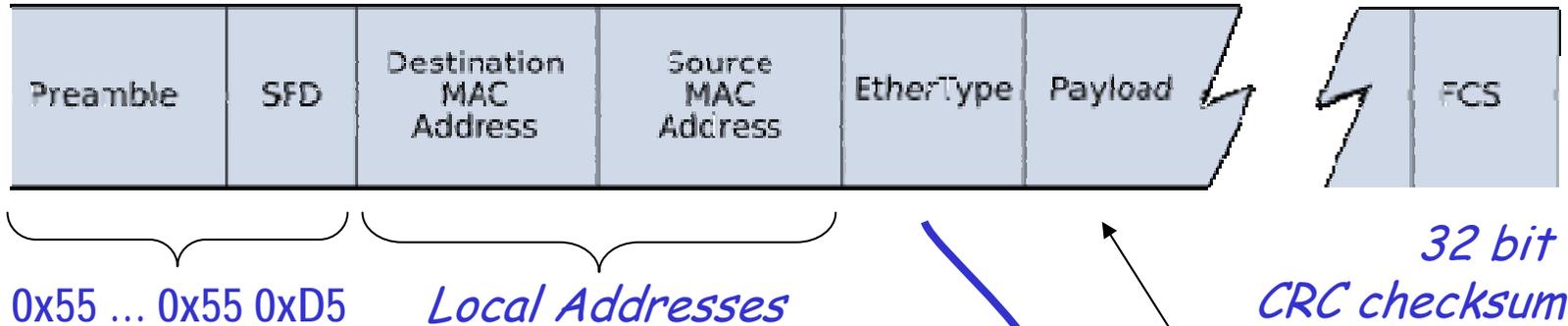
- Substandards:*
- 802.1: Zusammenhang der Standards und MAC Bridging
 - 802.2: Logical-Link-Control-Dienste/Protokolle (LLC)
 - 802.3: CSMA/CD-Protokoll auf Bustopologie
 - 802.4: Token-Bus-Protokoll auf Bustopologie
 - 802.5: Token-Ring-Protokoll auf Ringtopologie
 - 802.10: Sicherheitsstruktur für 802-Protokolle
 - 802.11: Wireless LANs
 - 802.14: Cable Modem
 - 802.15: Wireless Personal Area Networks
 - 802.16: Broadband Wireless Access

Addressing in 802-LANs



- IEEE 802 uses 48 bit addresses to identify the stations (both source and destination of a frame).
- Flags (2 bit) distinguish
 - Unicast and multicast, i.e. only one station addressed or a group of stations
 - Globally unique versus locally administered addresses
 Note: Locally administered addresses do not contain a organizationally unique identifier (OUI), but use a flat 46 bit space.
- Special address FF:FF:FF:FF:FF:FF addresses all stations (=broadcast)

Ethernet Frame



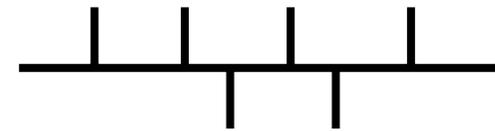
Typical data link layer frame headers have a field that indicates the network layer protocol that is carried in the respective frame.

RFC 1010, 1042, ... define

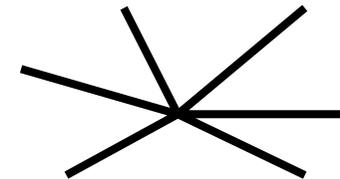
- 0x0800 Internet Protocol, Version 4 (IPv4)
- 0x0806 Address Resolution Protocol (ARP)
- 0x8035 Reverse Address Resolution Protocol (RARP)
- 0x86DD Internet Protocol, Version 6 (IPv6)

Logical versus Physical Structure

- Lokale Netze bilden häufig einen logischen Bus, obwohl eine physische Stern- oder Baumstruktur vorliegt
 - Ethernet: typisch als Stern verkabelt, aber logisch ein Bus
 - Universal Serial Bus: physische Baumstruktur, aber logisch ein Bus
- Logischer Bus:
Der Mittelpunkt eines physischen Sterns (Hub) kann nicht ausschließlich an die richtige Station weiterleiten. Folglich erhalten alle Stationen das gleiche Signal, als ob sie an einen Bus angeschlossen wären



Bus

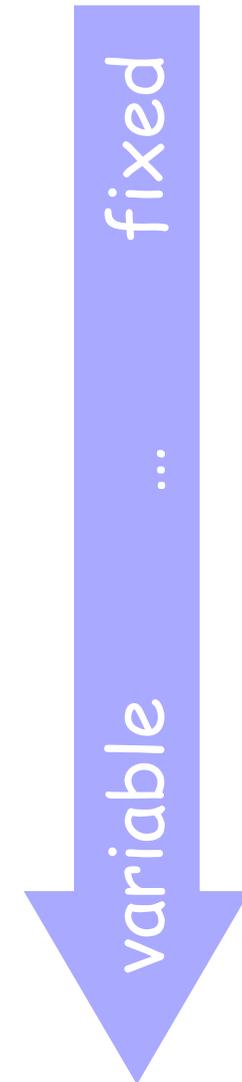


Star

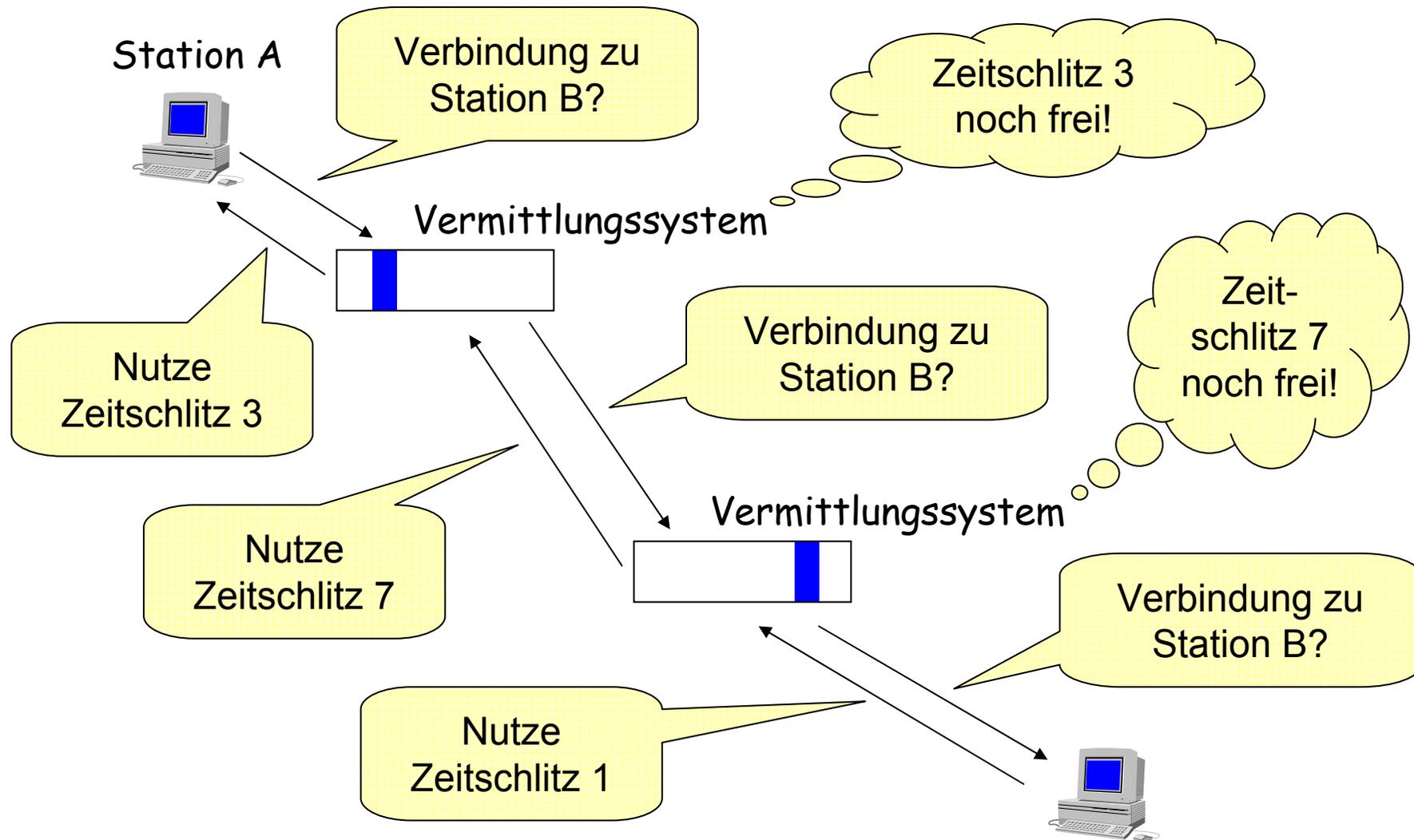


Medium Assignment Principles

- Circuit Switched Networks
 - Connection setup assigns fixed slots for all stations
 - Examples: ISDN, ...
- Master Slave with Polling
 - A special station (=master) polls all other stations (=slaves), e.g. in a round-robin manner
 - Example: Bluetooth
- Priorities
 - Physical properties achieve arbitration between stations
 - Example: CAN bus
- Token Based
 - Stations circulate a token.
 - Only token-holders may access the medium.
- Randomized Access Control
 - Examples: Aloha, CSMA/CD („Ethernet“), ...

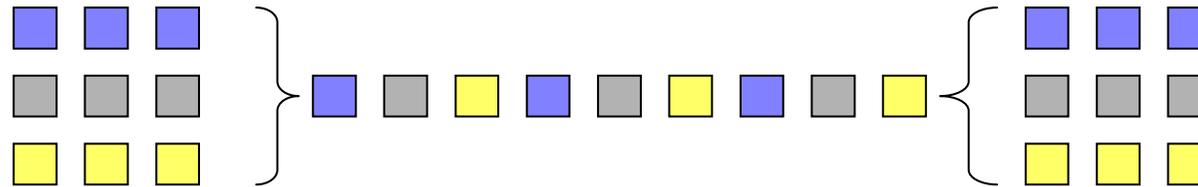


Feste Zuteilung bei der Leitungsvermittlung

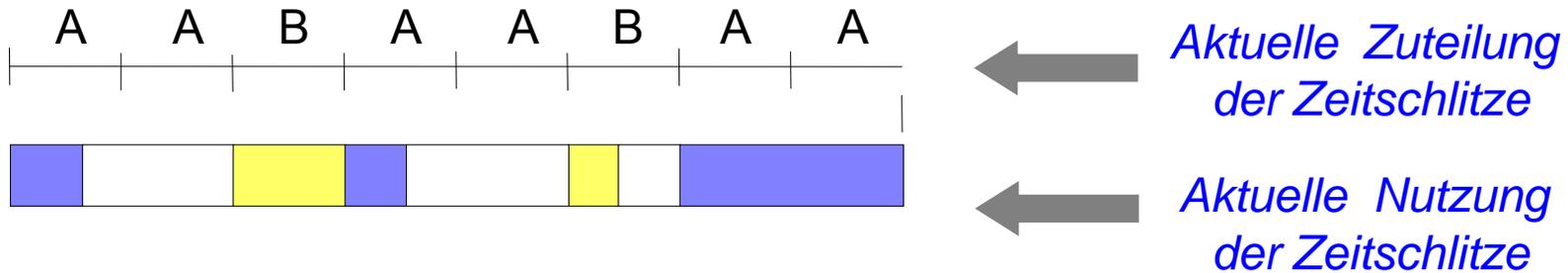


Feste Zuteilungsstrategie

Gleichmäßige Verteilung und Nutzung:



Ungleichmäßige Verteilung und Nutzung:



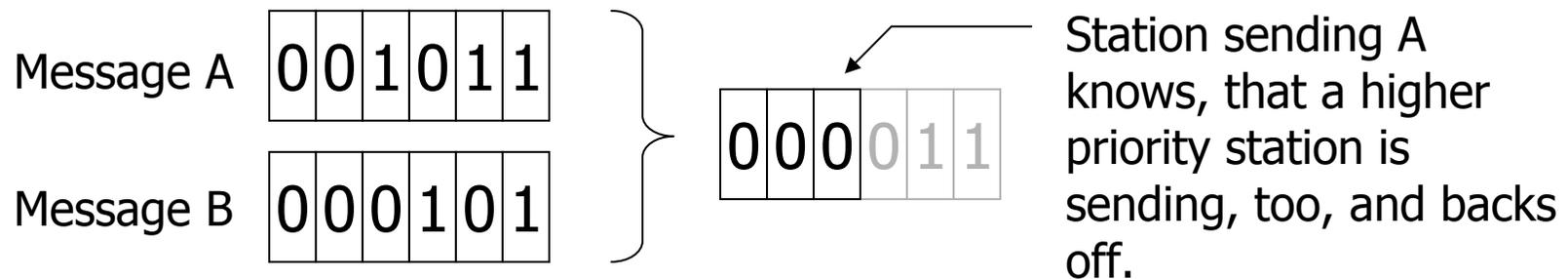
Benutzer A
 Benutzer B
 Ungenutzt

Feste Zuteilungsstrategie

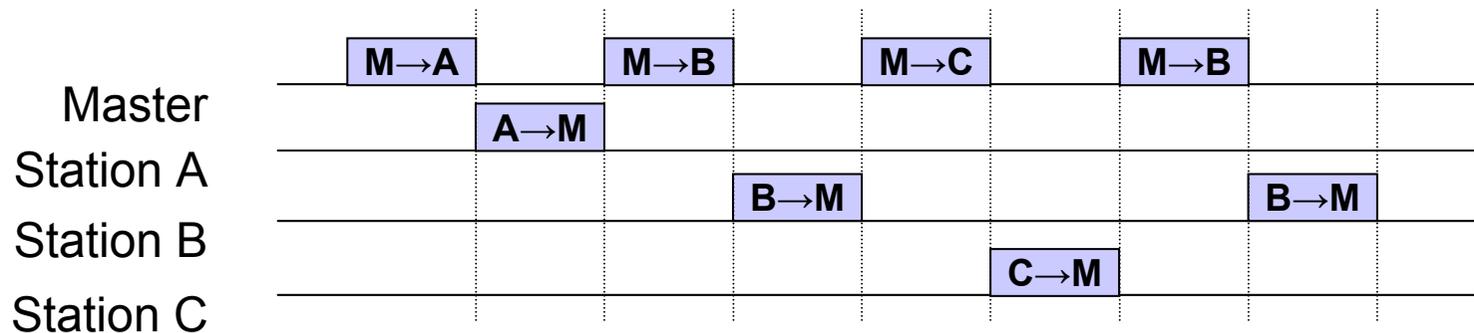
- Alle angeschlossenen Stationen bekommen eine sich periodisch wiederholende feste Redezeit zugewiesen.
- Eine Station kann mehrere Zeitschlitzze erhalten.
- Vorteile:
 - Einfach zu implementieren
 - Kaum Overhead während der Kommunikation
 - Rederecht für alle regelmäßig garantiert
 - Kein Jitter (=schwankende Verzögerungen)
- Nachteile:
 - Synchronisierung erforderlich
 - Hinzufügen von Geräten schwierig (Vergabe der Zeitschlitzze muss geregelt werden → Leitungsvermittlung)
 - Datenrate nur in Stufen variierbar, d.h. schwankende Anforderungen nicht unterstützt
 - Kapazität des Kanals kann nicht voll genutzt werden, falls Geräte unterschiedliche bzw. schwankende Datenraten haben
 - Ggf. höhere Verzögerungen

Other Medium Access Control Mechanisms

CAN (=Controller area network) bus uses arbitration:



Bluetooth uses time slots and polling by a master:



Controller Area Network („CAN-Bus“)

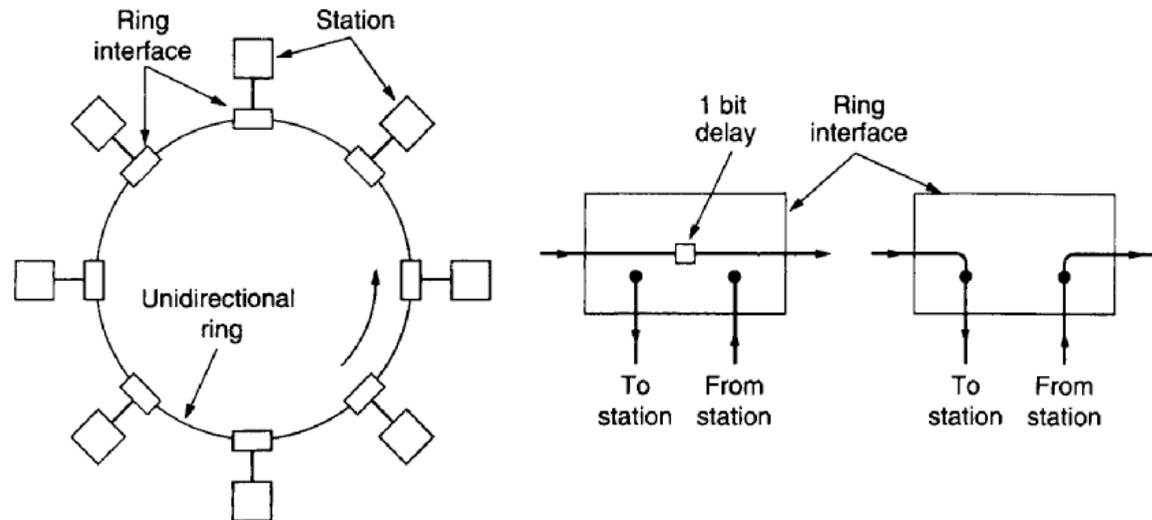
- Entwickelt von Bosch (ISO-Standard 11898)
 - Kostengünstige, zuverlässige Feldbus-Lösung für zeitkritische Anwendungen
 - Anwendung: Vernetzung von Sensoren und Aktoren in Fahrzeugen und anderen Maschinen
- Eigenschaften
 - Datenrate 10kBit/s bis 1MBit/s
 - Netzausdehnung 40m bis 5000m
 - Stationen: unbegrenzt
 - Medien
 - Bus mit verdrehter Doppelader (Twisted Pair) und 120Ω Abschlußwiderstand
 - Glasfaser mit Optokopler (als logischer Bus organisiert)
- Nachrichten als Rahmen mit variabler Länge (bis zu 8 Byte Nutzdaten)
- Adressierung von Inhalten, nicht Stationen
 - Unbegrenzte Zahl von Stationen
 - Begrenzte Zahl von Nachrichten-Typen (Adressen mit 11 Bit, erweitert 29 Bit)
- Adressierung mittels Identifier
 - Stationen können Nachrichten filtern
 - So können auch mehrere Stationen gleichzeitig eine Nachricht bearbeiten (=Multicast)
 - Identifier haben unterschiedliche Prioritäten (z.B. kleinste Nummer hat höchste Priorität, d.h. Nachrichten werden in aufsteigender Reihenfolge der Identifier versandt)

Bluetooth – Master-Slave w/ Polling

- Alternative Variante von festen Zeitschlitzten: Master-Slave mit Polling
- Unter den angeschlossenen Geräten übernimmt eines die Rolle des Masters und regelt die Kommunikation, in dem es das Rederecht zeitlich begrenzt an die Slaves vergibt.
- Der Master kann bestimmt sein durch
 - Besonderer Gerätetyp
 - Anschlussart, z.B. Wurzelknoten in einer Baumstruktur
 - Zufälliges Auswahlverfahren
- Ist in einem System ein Master identifiziert, kann dieser alle Stationen der Reihe nach abfragen
 - Slaves senden nur nachdem sie vom Master angesprochen wurden
 - Master bestimmt, wie oft (und wie lange) Slaves senden dürfen
 - Beispiel: Bluetooth
- Bewertung
 - Einfaches Verfahren falls ein Master bereits festgelegt ist
 - Wahl eines Masters aus gleichartigen Geräten schwierig
 - Ausfall des Masters kann Kommunikation völlig verhindern
 - Kommunikationsbeginn durch Slaves schwierig

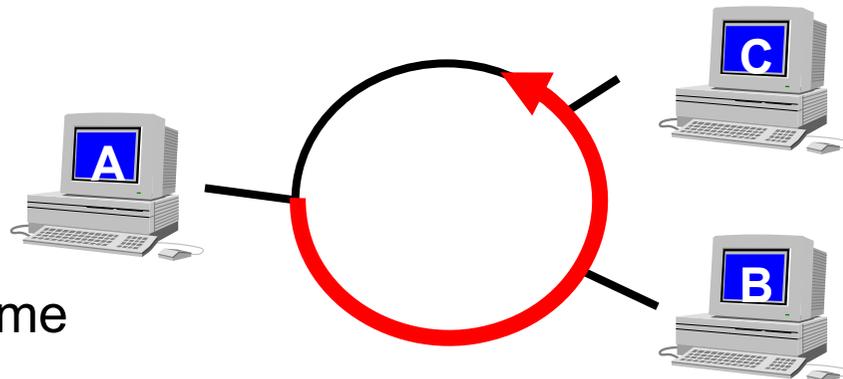
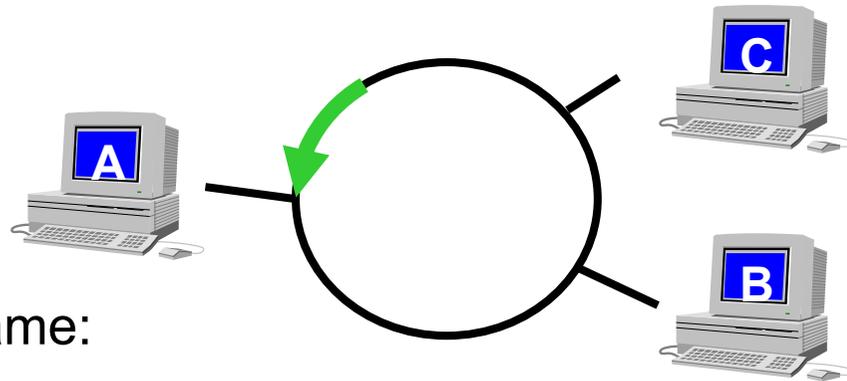
Token-Ring

- Alle Stationen sind als Ring verbunden
- Datenstrom wird jeweils weitergereicht
- Ist das Medium frei, kann ein Sender einen Rahmen „auf den Ring legen“
- Er muss diesen Rahmen nach einem Umlauf wieder entfernen



Token Ring – Token Based MAC

- Stations forward bitstream immediately to their neighbor.
- Ring contains a special frame: the token
- If a station wants to send, it replaces the token by its frame of data.
- The receiver copies the data, but forwards the frame, too.
- The sender replaces its frame with the token.

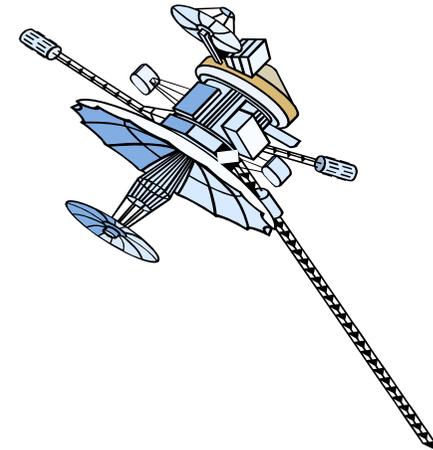


Randomized Access Control

- Alternative Variante von Zeitschlitz:
 - Zeitschlitz werden nicht fest zugewiesen
 - Ein Gerät sendet bei Bedarf
 - Kollision (d.h. gleichzeitiges Senden zweier oder mehrerer Stationen) wird in Kauf genommen
 - Dadurch Totalverlust der Rahmen in Zeitschlitz mit Kollision
- Bewertung
 - Besonders geeignet z.B. für Satellitennetze („Slotted Aloha“ – bereits in 70er Jahren verwendet im Uni-Netz in Hawaii)
 - Hinzufügen von Geräten einfach
 - Synchronisierung erforderlich
 - Effizienz nimmt bei steigendem Verkehr stark ab (Kollisionswahrscheinlichkeit steigt an)

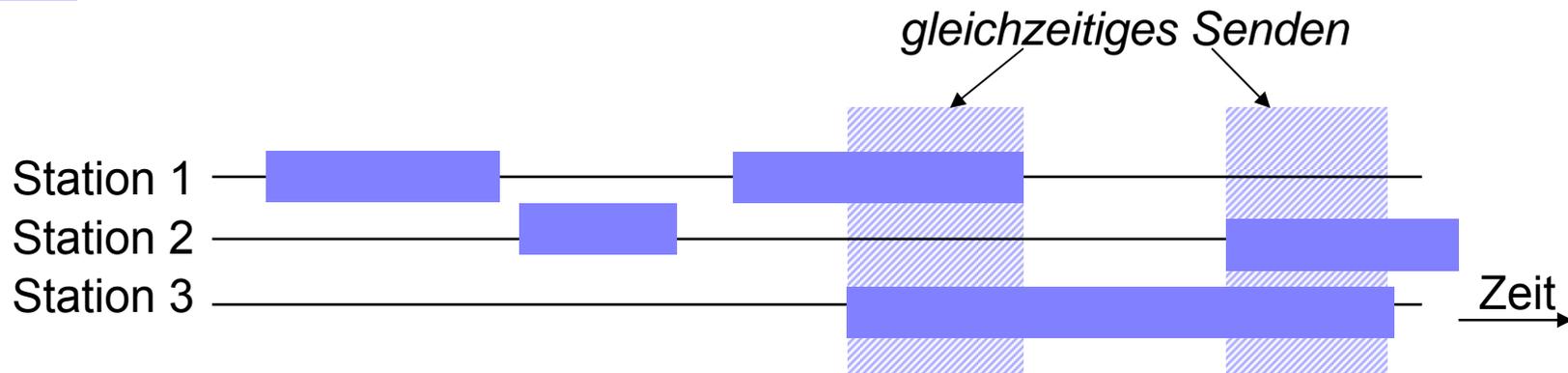
Beispiel: Aloha & Slotted Aloha

- Rechnernetz für die Hawaii-Inseln
- Satellitenfunksystem, d.h. Rundspruch-Medium, aber sehr lange Laufzeit
- Lösung 1: Aloha
 - Jeder sendet wann er will.
 - Bei einer Kollision sind beide Rahmen zerstört.
- Lösung 2: Slotted Aloha
 - Wie oben, aber Zeitschlitzte fester Länge.
 - Deren Synchronisation kann vorausgesetzt werden.
 - Stationen senden in beliebigen Zeitschlitzten.
- Slotted Aloha ist doppelt so effizient, da Kollisionen innerhalb der Zeitschlitzte begrenzt bleiben.

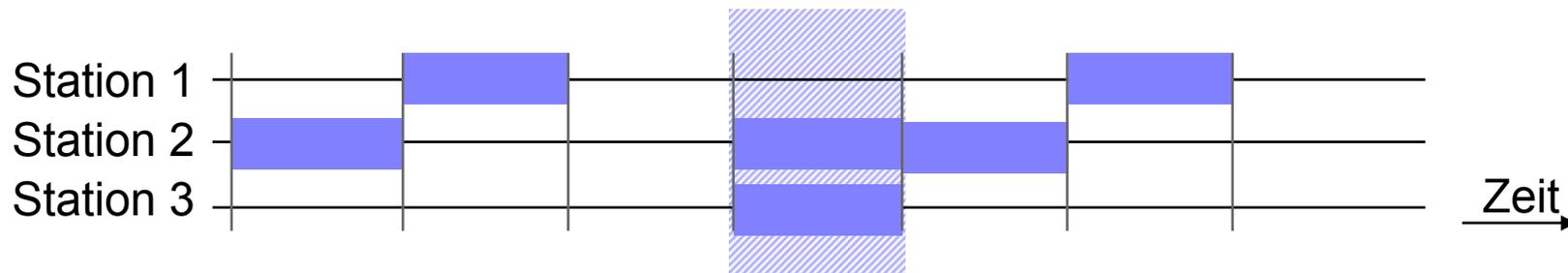


Beispiel: Aloha & Slotted Aloha

Aloha:



Slotted Aloha:



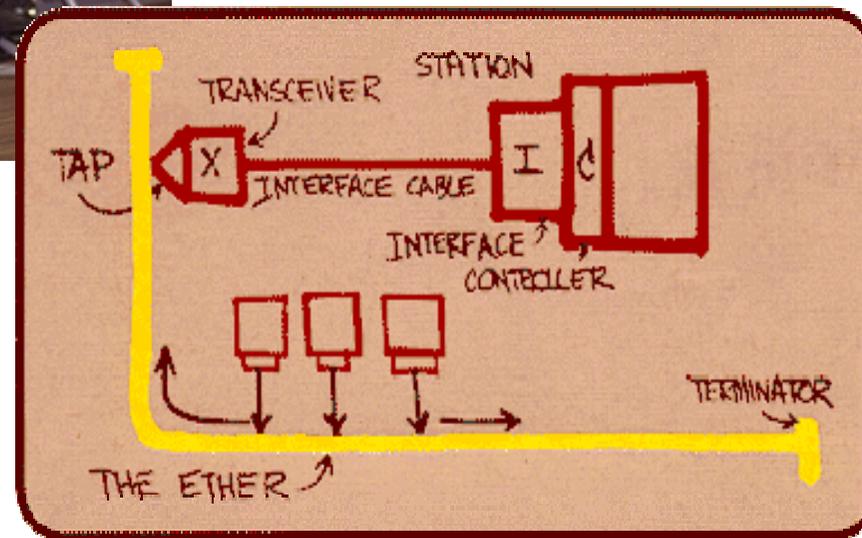


Bob Metcalfe entwickelt Ethernet 1973-76 bei Xerox

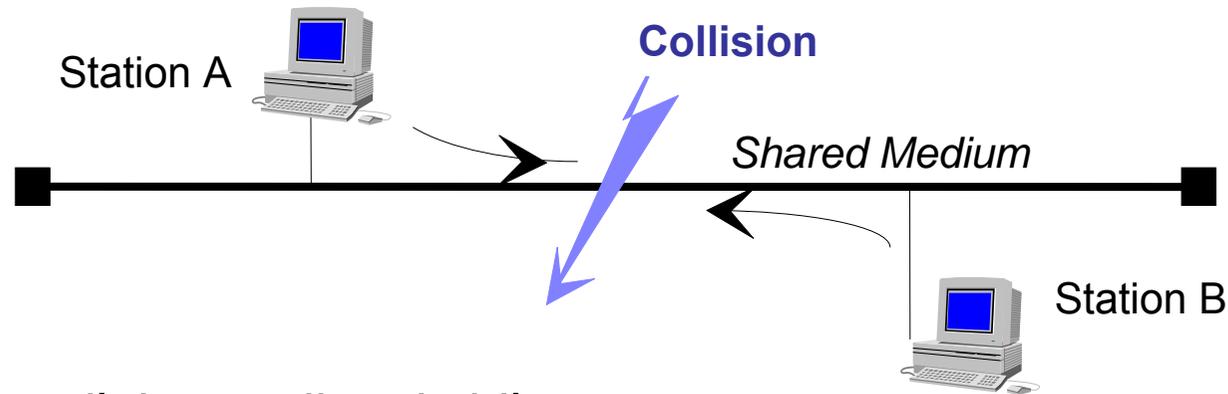
- Physischer Bus
- CSMA/CD
- 2.94MBit/s

Heute

- Meist Sternverkabelung
- Aktive Zwischensysteme: Switches
- Bis zu 10GBit/s



Ethernet – MAC using CSMA/CD



- Wait until the medium is idle.
- Start sending ...
- If a collision is detected, don't continue with the frame but send a special signal that makes all stations discard the data.
- Wait some time* and retransmit the frame.

* Different waiting strategies lead to different protocol flavors.

Beispielablauf

A beginnt Übertragung (t_0)



B beginnt Übertragung
($t_0 + t_L - \epsilon$)



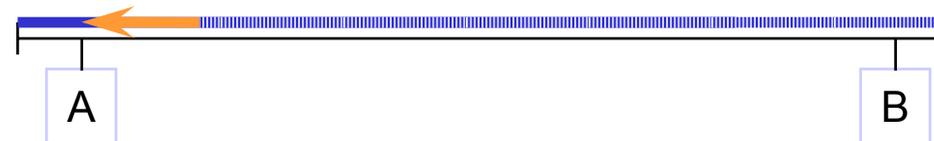
B entdeckt Kollision, stoppt
eigene Übertragung ($t_0 + t_L$)



B schickt Jamming-Signal,
Kollision läuft weiter



A entdeckt Kollision
($t_0 + 2(t_L - \epsilon)$)



t_L Signallaufzeit von A nach B (Propagation Delay)

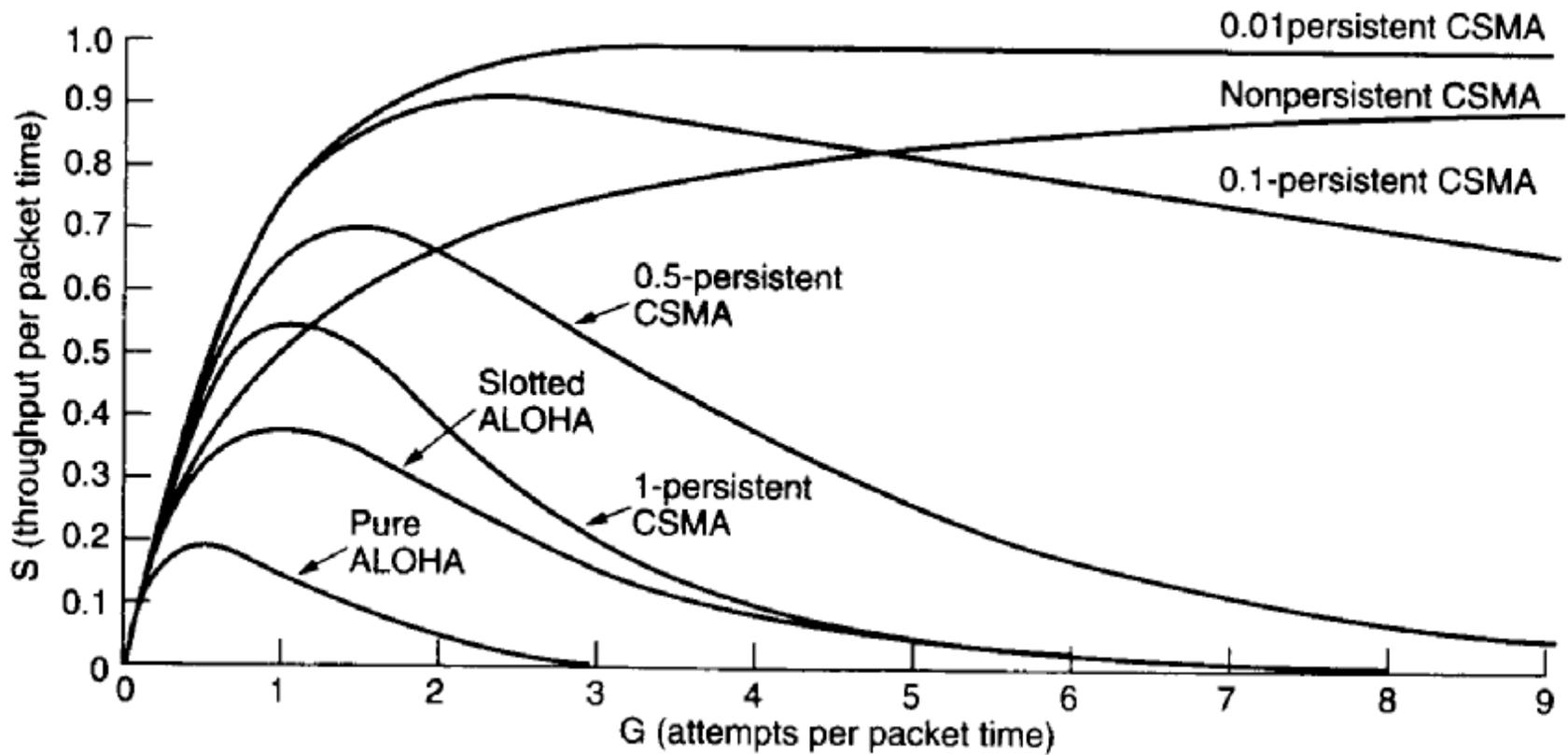
$2 t_L$ Signallaufzeit von A nach B und zurück (Round Trip Delay)

- Wie wird nach einem erfolglosen Medienzugriff, d.h. das Medium war belegt, verfahren? – Mögliche Strategien
 - Nicht persistent
 - Ist der Kanal frei, so wird sofort wieder gesendet.
 - Ist der Kanal belegt, so wartet die Station eine zufällig berechnete Zeitspanne und startet einen neuen Versuch für den Medienzugriff.
 - 1-persistent
 - Station hört weiterhin das Medium ab und startet den eigenen Medienzugriff sofort nach Abschluss der laufenden Übertragung (Sendewahrscheinlichkeit = 1)
 - p-persistent ($0 < p < 1$)
 - Station hört weiterhin das Medium ab. Nach Ende der laufenden Übertragung sendet sie ihre Daten mit einer vorher festgelegten Wahrscheinlichkeit p oder wartet mit Wahrscheinlichkeit $1 - p$ eine festgelegte Zeitspanne t . Falls das Medium bei der ersten Überprüfung frei ist, wird ebenso verfahren.
- Implementierung muss einen Kompromiss finden zwischen
 - der Minimierung der Anzahl von Kollisionen und
 - einer möglichst permanenten Nutzung des Übertragungsmediums
- Bei Ethernet wird eine 1-persistente Strategie verwendet.
 - Welche Probleme können dadurch entstehen?

- Alle Stationen überwachen das Medium und senden wenn es frei ist bzw. eine gewisse Zeit frei war
 - *Carrier Sense Multiple Access* (CSMA) Protokoll
 - Zufällige Wahl der Wartezeit ermöglicht, dass meist nur eine Station sendet (*Non-persistent CSMA*)
 - Gleichzeitiger Sendestart verhindert „Verschnitt“, erhöht aber Kollisionswahrscheinlichkeit (*1-persistent CSMA*)
- Bewertung
 - Einfaches Verfahren, robust gegen Ausfälle von Stationen
 - Flexible Handhabung vieler Stationen, da keine festen Zuweisungen erforderlich
 - Keinerlei Garantien, dass Übertragung erfolgreich sein kann
 - Kapazität des Kanals kann nicht voll genutzt werden

- Das Medienüberwachungsverfahren kann dadurch verbessert werden, dass Stationen eine Kollision selbst erkennen
 - Bei einer Kollision brechen beide Stationen ab und wiederholen die Übertragung nach einer zufälligen Zeit
 - Es muss sichergestellt sein, dass Stationen sich einig sind, ob eine Kollision vorlag oder nicht (Mindestlänge für Rahmen, Jamming-Signal)
- Dieses Verfahren (*Carrier Sense Multiple Access with Collision Detect = CSMA/CD*) wird u.a. bei Ethernet verwendet

Effizienz der CSMA-Protokolle



Data Link Layer – Summary

1. Packet serialization and de-serialization (=Framing)



2. Deal with bit errors (=Checksums and Forward Error Correction)

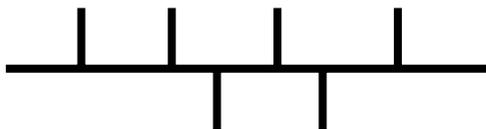
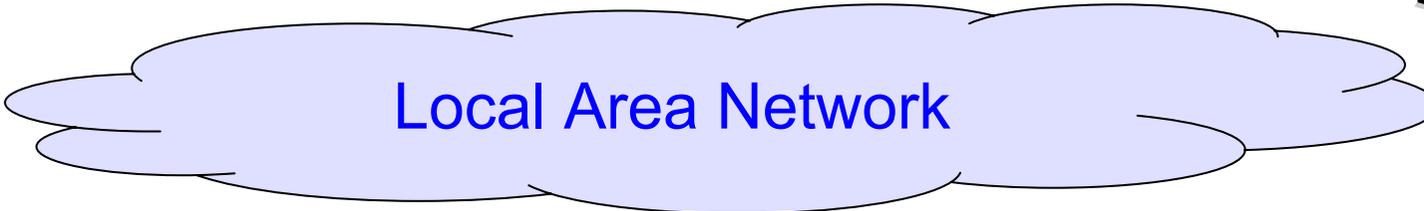
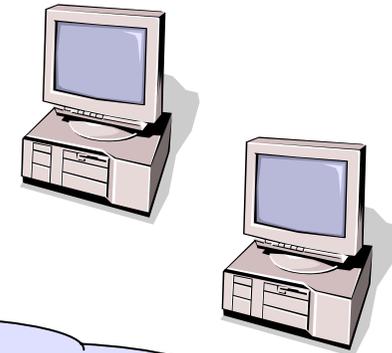


3. Contention in a shared medium (→Ethernet, Token Ring, etc.)

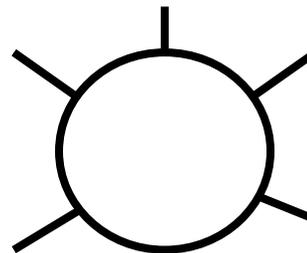


Data Link Layer

Stations can send datagrams to any or all other station(s) in the same LAN. Some LAN technologies provide a reliable service, others don't.



Bus e.g. Ethernet

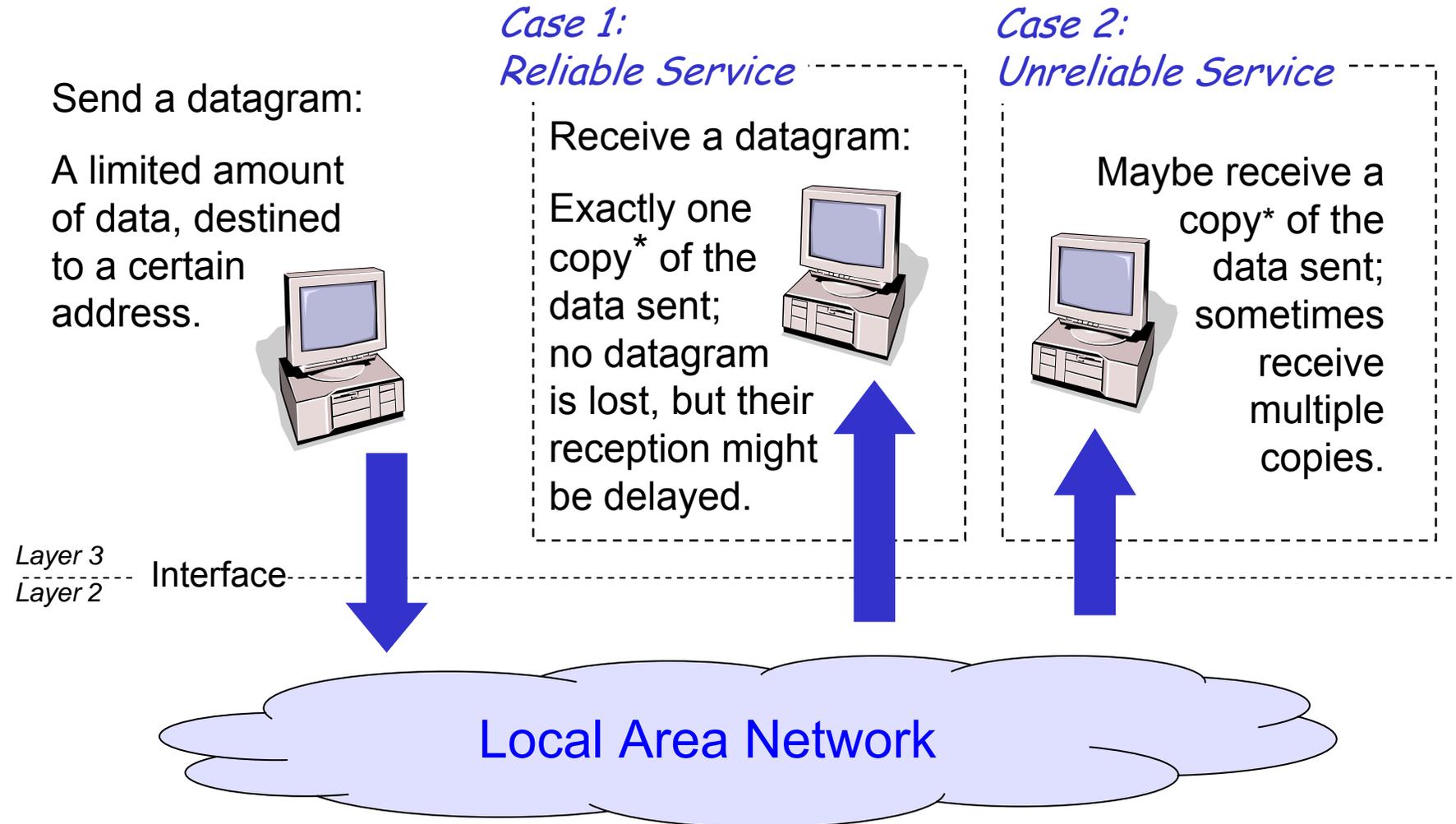


Ring e.g. Token Ring

etc.

e.g. Wireless LAN

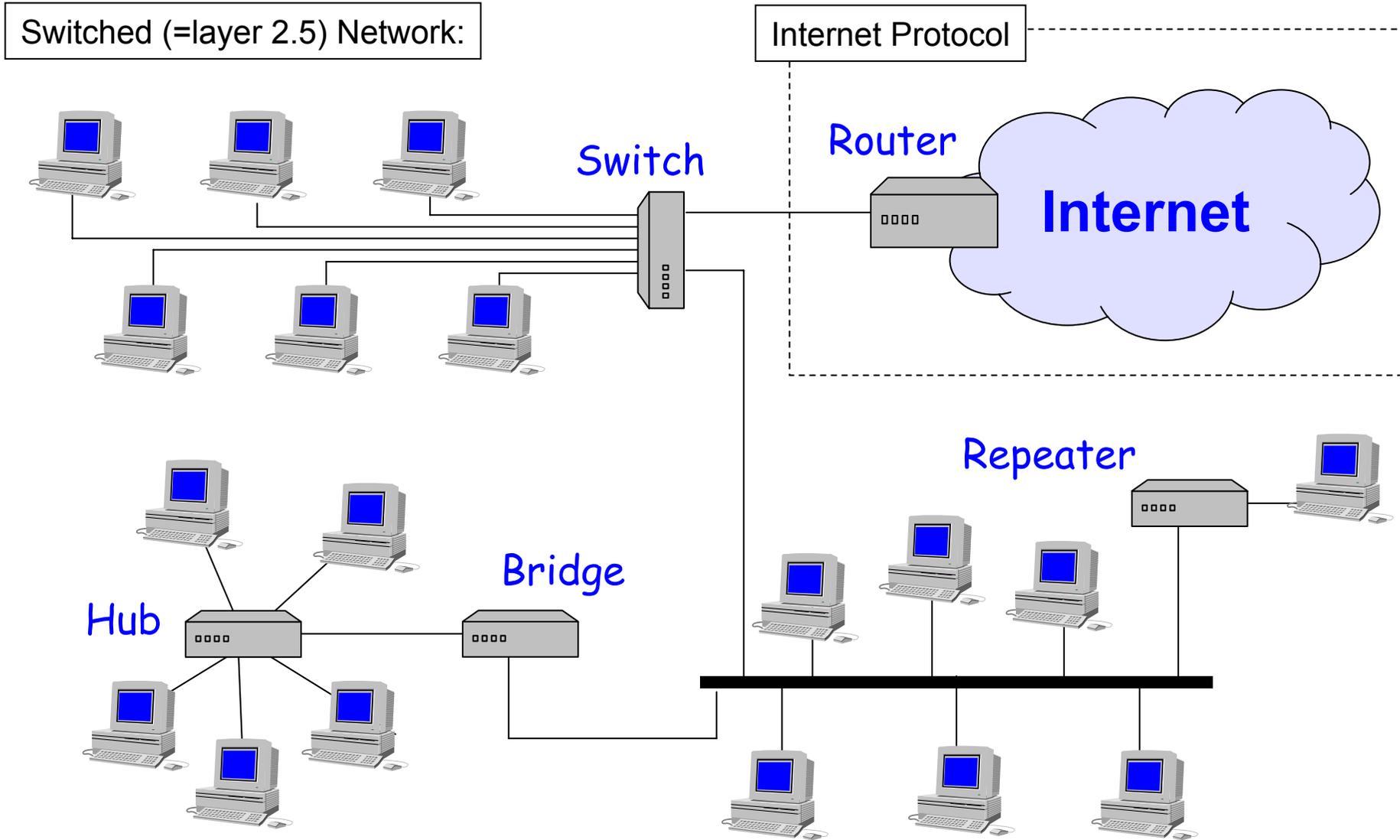
Data Link Layer



* Only in rare cases, the data might be corrupted.

- The data link layer is concerned with the local delivery of datagrams:
 - Physically direct connected stations (same cable, etc.)
 - Logically / physically close stations, e.g. in the LAN of an office building, campus, etc.
 - Logically close, e.g. in the network of the same provider
- Depending on the physical directness / indirectness:
 - Analog repeater: Amplify the signal, but do not touch the modulation.
 - Digital repeater: Amplify, reshape and retime the modulated signal.
 - Hub: Physically connect different network segments (=typically cables)
 - Bridge: Logically connect different network segments
 - Switch: Connect different network segments and perform more advanced protocol interactions such as forwarding datagrams only into some of the segments.

Layer 2 / 2.5 Networks



Repeater, Bridges und Switches

- Ein LAN (Lokal Area Network) kann aus verschiedenen Medienabschnitten bestehen, z.B. unterschiedliche Kabelstränge auf einzelnen Stockwerken
 - *Repeater* reicht Bitstrom weiter (Schicht 1)
 - *Bridge* (Brücke) reicht Rahmen weiter (Schicht 2)
Transparente Brücken verhalten sich so, als ob die angeschlossenen LANs ein einziges wären
 - *Switch* reicht Rahmen möglichst nur an die jeweiligen Empfänger weiter (Schicht 2 bzw. 3)
- Bridges und Switches können LANs mit verschiedenen Protokollen verbinden (z.B. Ethernet und Wave-LAN, d.h. Funknetze)

Switched versus Routed Networks

- In theory, switched networks can become very large.
- Often, the network of an entire organization (e.g. global company network, University network, etc.) is switched.
- The Internet Protocol can
 - ... route datagrams across different layer 2 networks
 - ... more efficiently maintain routing information in very large networks
 - ... more efficiently restrict traffic to those network parts that need that traffic (→ i.e. avoids broadcast, see later)

- Forwarding
 - A brief review of packet switched networks
 - Discussion of store & forward
 - Local versus global addresses (→ARP)
- Routing in the Internet
 - Static and dynamic protocols – An overview
 - Bellman-Ford and Dijkstra (→RIP & OSPF)
 - Millions of hosts – The quest for scalability
- The Internet Protocol (version 4)
 - The header fields' who is who

Why does data link layer switching not work globally?

What's better in circuit switched networks than in the Internet?

Questions?



Thomas Fuhrmann

CS VIII - Network Architectures
Technical University Munich, Germany

IBDS System Architecture
University of Karlsruhe, Germany

fuhrmann@net.in.tum.de