



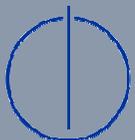
Lehrstuhl für Netzarchitekturen und Netzdienste
Institut für Informatik – Technische Universität München
Prof. Dr.-Ing. Georg Carle

Grundlagen: Rechnernetze und Verteilte Systeme

Kapitel 4: Vermittlung

Paket-/Leistungsvermittlung, Brücke, Router

Prof. Dr.-Ing. Georg Carle
Lehrstuhl für Netzarchitekturen und Netzdienste
Technische Universität München
carle@net.in.tum.de
<http://www.net.in.tum.de>





Motivierende Fragen

- ❑ Was versteht man unter Repeater und Brücken ?
- ❑ Wie funktioniert ein LAN ?
- ❑ Was versteht man unter einer strukturierten Verkabelung ?
- ❑ Welche Arten von Vermittlungen in globalen Netzen gibt es ?
- ❑ Was ist ein Router und wie funktioniert er ?



Übersicht

1. Einführung und Motivation
 - Bedeutung, Beispiele
2. Begriffswelt und Standards
 - Dienst, Protokoll, Standardisierung
3. Direktverbindungsnetze
 - Fehlererkennung, Protokolle
 - Ethernet
4. **Vermittlung**
 - **Vermittlungsprinzipien**
 - **Wegwahlverfahren**
5. Internet-Protokolle
 - IP, ARP, DHCP, ICMP
 - Routing-Protokolle
6. Transportprotokolle
 - UDP, TCP
7. Verkehrssteuerung
 - Kriterien, Mechanismen
 - Verkehrssteuerung im Internet
8. Anwendungsorientierte Protokolle und Mechanismen
 - Netzmanagement
 - DNS, SMTP, HTTP
9. Verteilte Systeme
 - Middleware
 - RPC, RMI
 - Web Services
10. Netzsicherheit
 - Kryptographische Mechanismen und Dienste
 - Protokolle mit sicheren Diensten: IPSec etc.
 - Firewalls, Intrusion Detection
11. Nachrichtentechnik
 - Daten, Signal, Medien, Physik
12. Bitübertragungsschicht
 - Codierung
 - Modems



Ziele

- In diesem Kapitel wollen wir vermitteln
 - Grundverständnis von Netzwerkkoppelung
 - Funktionalität von Repeater, Brücken und LANs
 - Vermittlungsprinzipien
 - Funktionalität von Routern



Kapitelgliederung

4.1. Netzwerkkopplung

4.1.1. Repeater

4.1.2. Hub

4.1.3. Brücke (Bridge)

4.1.4. Spanning-Tree-Algorithmus

4.1.5. Remote-Brücke

4.1.6. Switched LAN

4.1.7. Virtuelle LANs

4.1.8. Leitbeispiel: Strukturierte Verkabelung

4.2. Vermittlungsprinzipien für globale Netze

4.2.1. Durchschaltevermittlung

4.2.2. Nachrichten-/Speichervermittlung

4.2.3. Paketvermittlung

4.2.4. Router

4.2.5. Routing-Verfahren

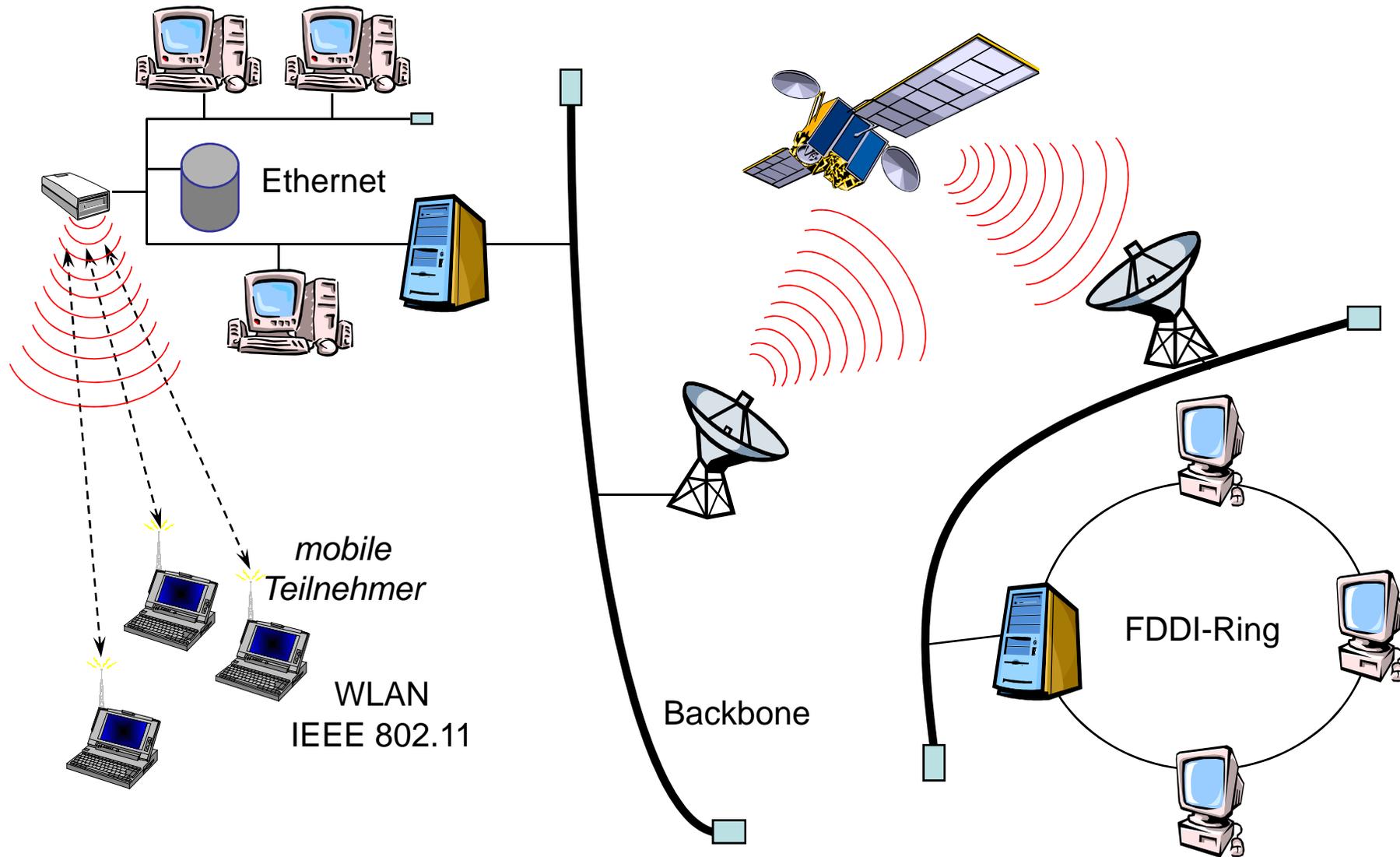


4.1. Netzwerkkopplung auf Schicht 2 (Ethernet)

- ❑ Repeater, Hub
- ❑ Brücke (Bridge)
- ❑ Spanning-Tree-Algorithmus
- ❑ Remote-Brücke
- ❑ LAN-Switch
- ❑ Virtuelle LANs
- ❑ Strukturierte Verkabelung

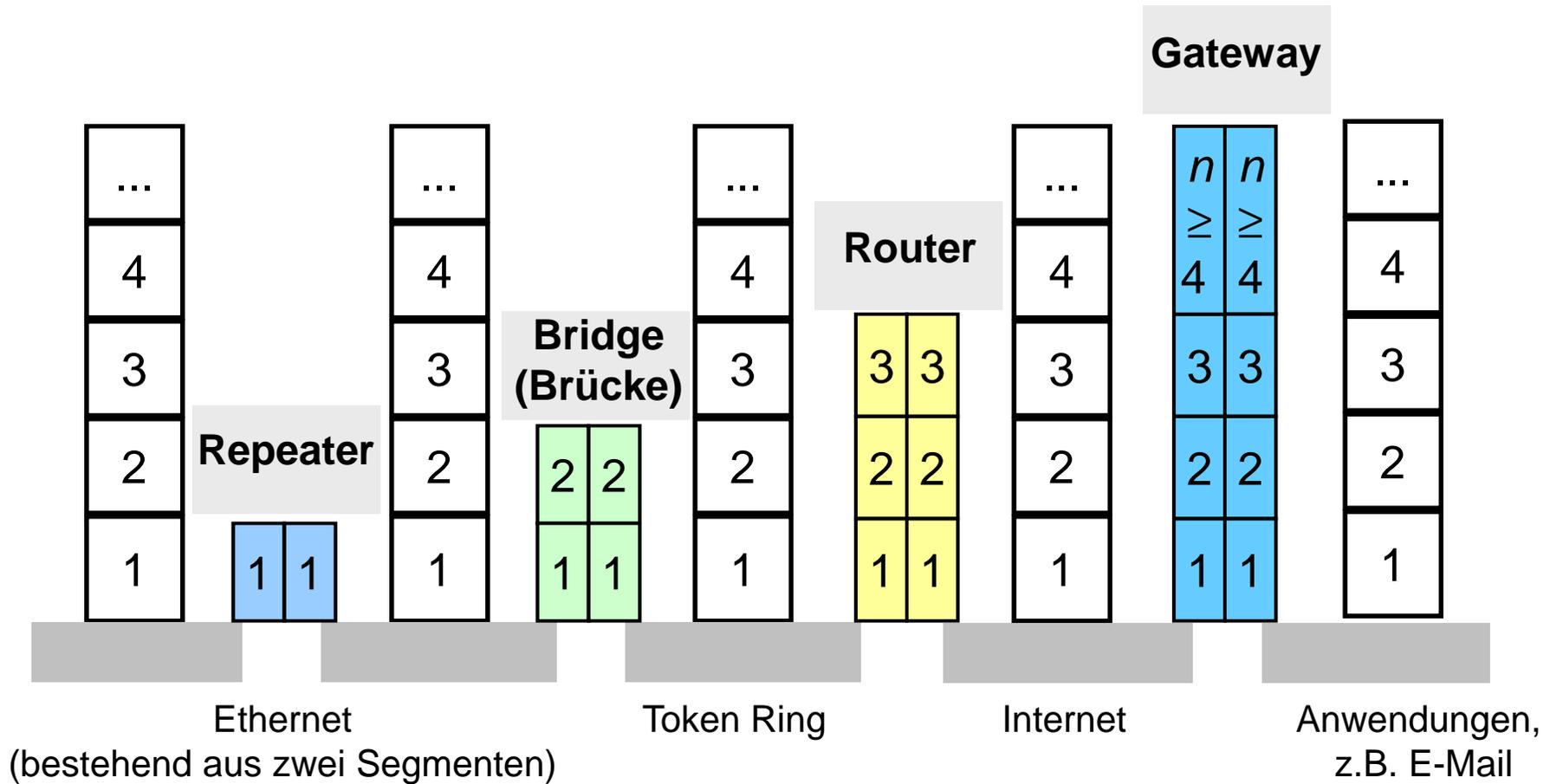


Netzwerkkopplung (allgemein)





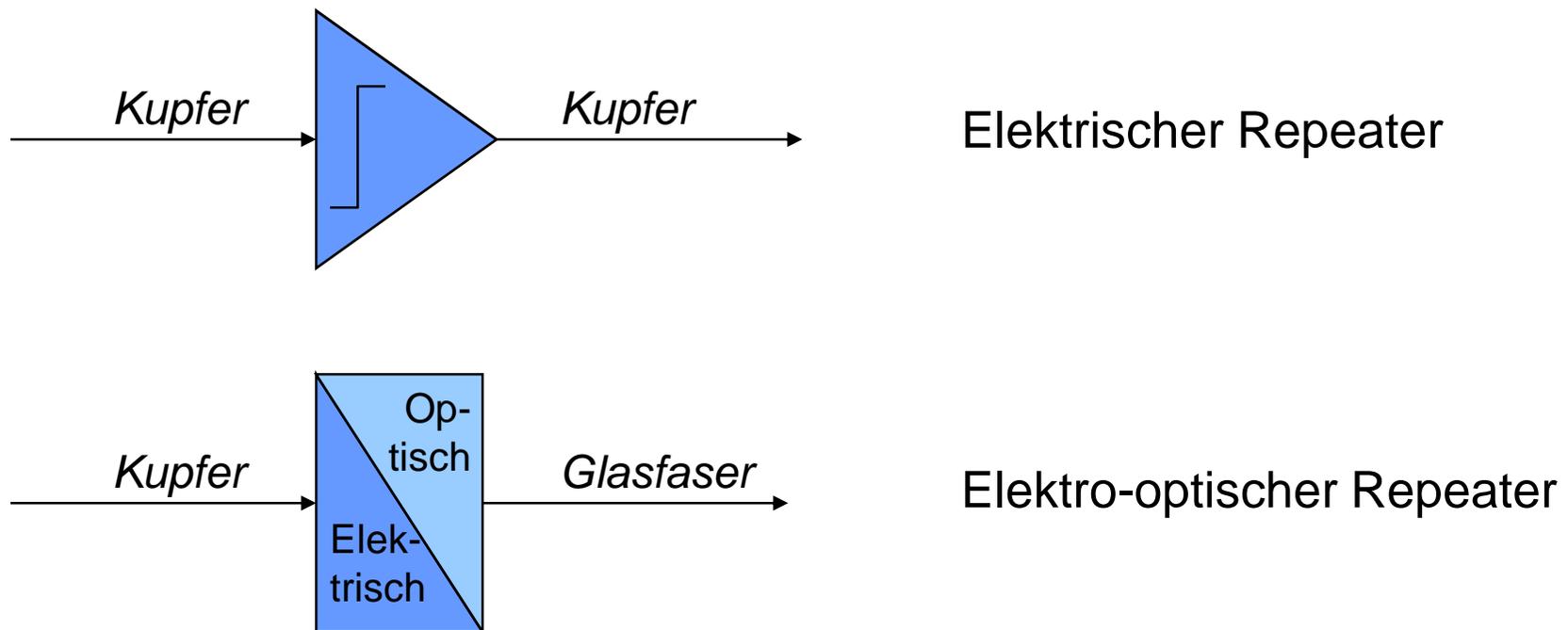
Kopplung von Netzen – Internetworking





Repeater

- ❑ Kopplung physikalischer Medien durch Signalregeneration/-verstärkung
- ❑ Keine Zwischenspeicherung
- ❑ Keine Bearbeitung der Pakete
- ❑ Medien können unterschiedlich sein, aber Protokoll auf Schicht 2 muss identisch sein





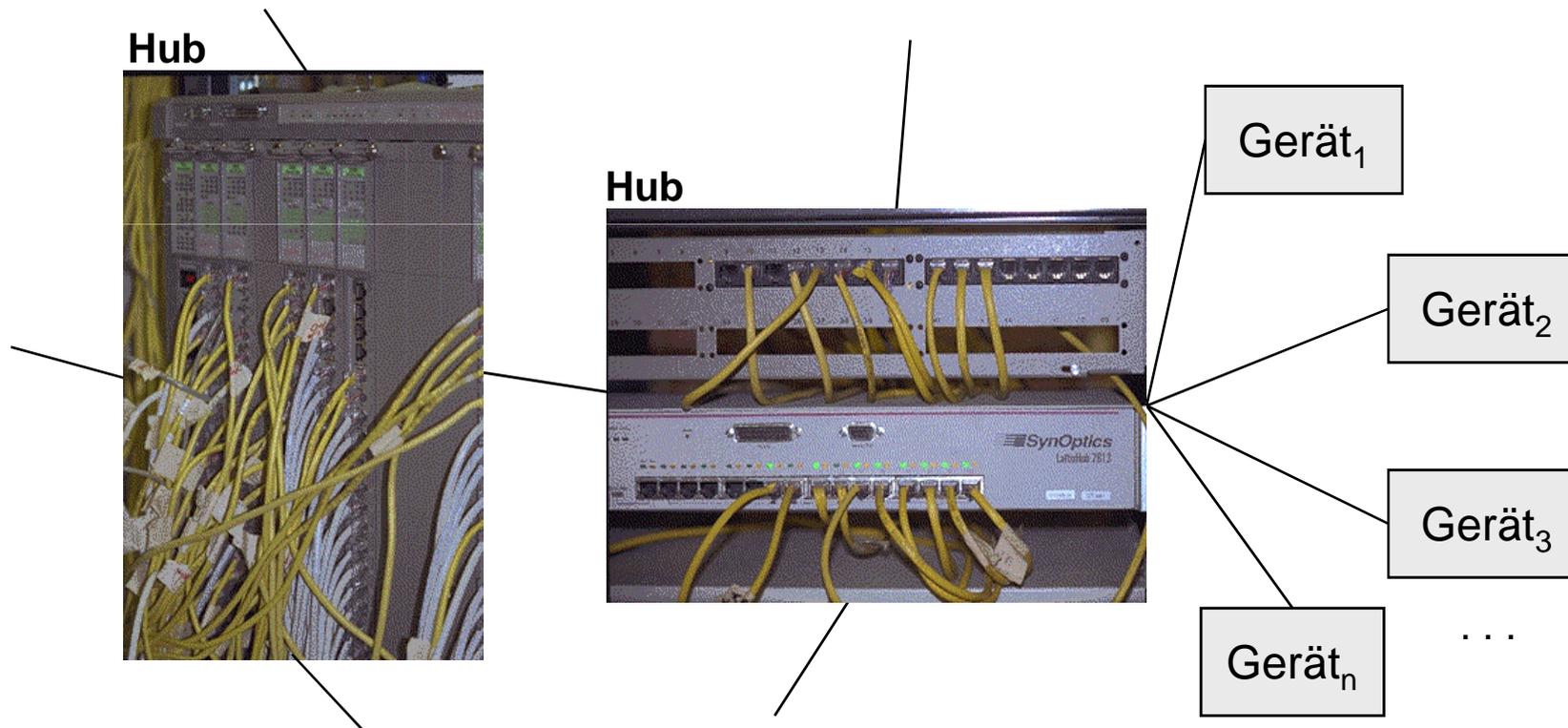
4.1.1. Repeater

- Kopplung von Netzwerken auf Schicht 1
 - Einsatzbereich:
 - Verbindung von lokalen Netzen zur Erhöhung der räumlichen Ausdehnung
 - Generierung mehrerer abgehender Signale an Verzweigungspunkten
 - Wechsel des Übertragungsmediums (z.B. Kupfer auf Glasfaser, Leitung auf Funkstrecke)
 - Vorteile:
 - Einfache Technik
 - Kostengünstige Lösung
 - Keine Verarbeitung an den Daten, somit keine Beeinträchtigung der Geschwindigkeit
 - Extrem lange Netzwerkverbindungen sind möglich (z.B. Überseeleitungen)
 - Nachteile:
 - Keine Intelligenz; alle Daten werden weitergeleitet
 - ⇒ **Keine Erhöhung der Netzkapazität durch Partitionierung**



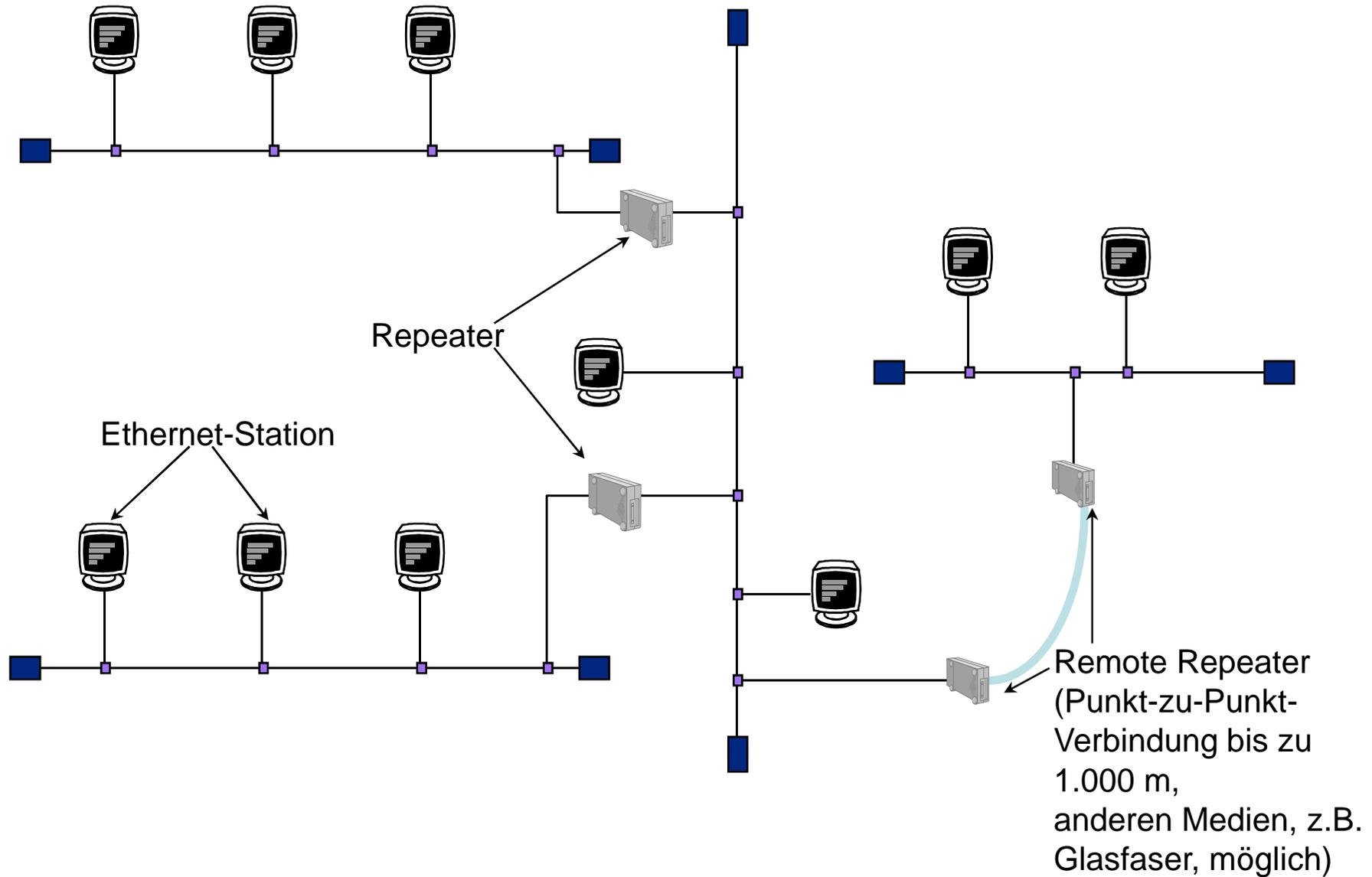
4.1.2. Hub

- Gleiche Funktion wie Repeater
 - typischerweise in Sterntopologie (separate Leitung von jedem Gerät zum Hub), kaskadierbar
 - Gesamtdurchsatz des Netzes wird nicht erhöht (im Gegensatz zu Switch)





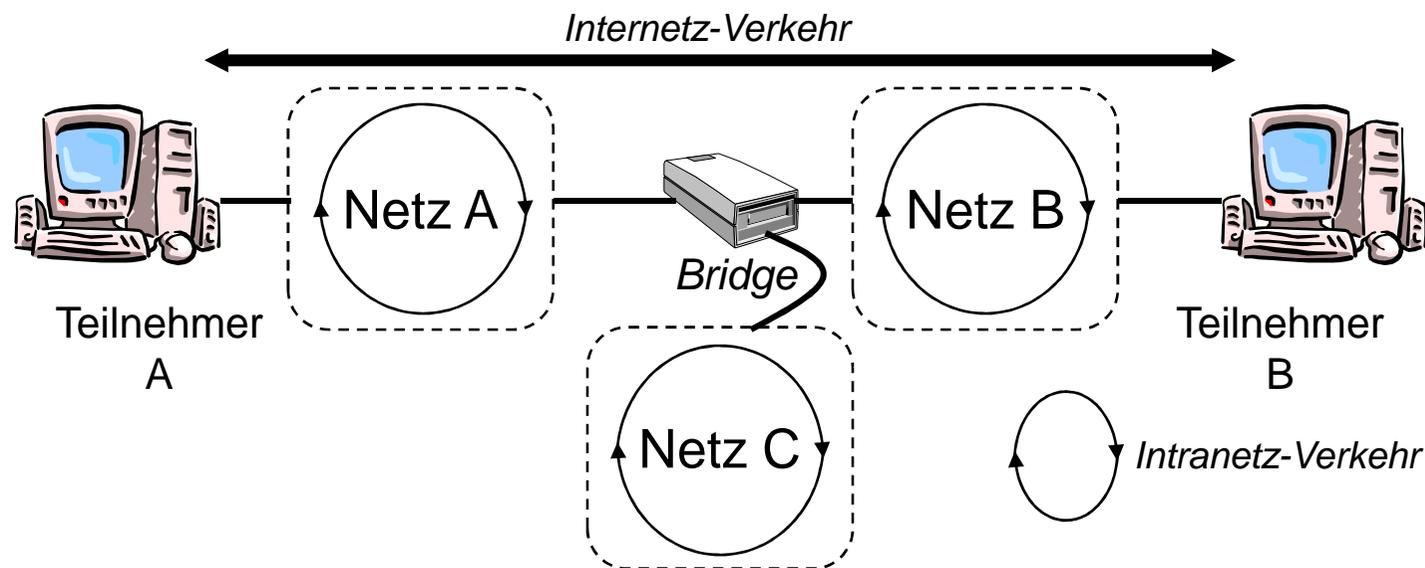
Netz-Kopplung auf Schicht 1: Beispiel Ethernet





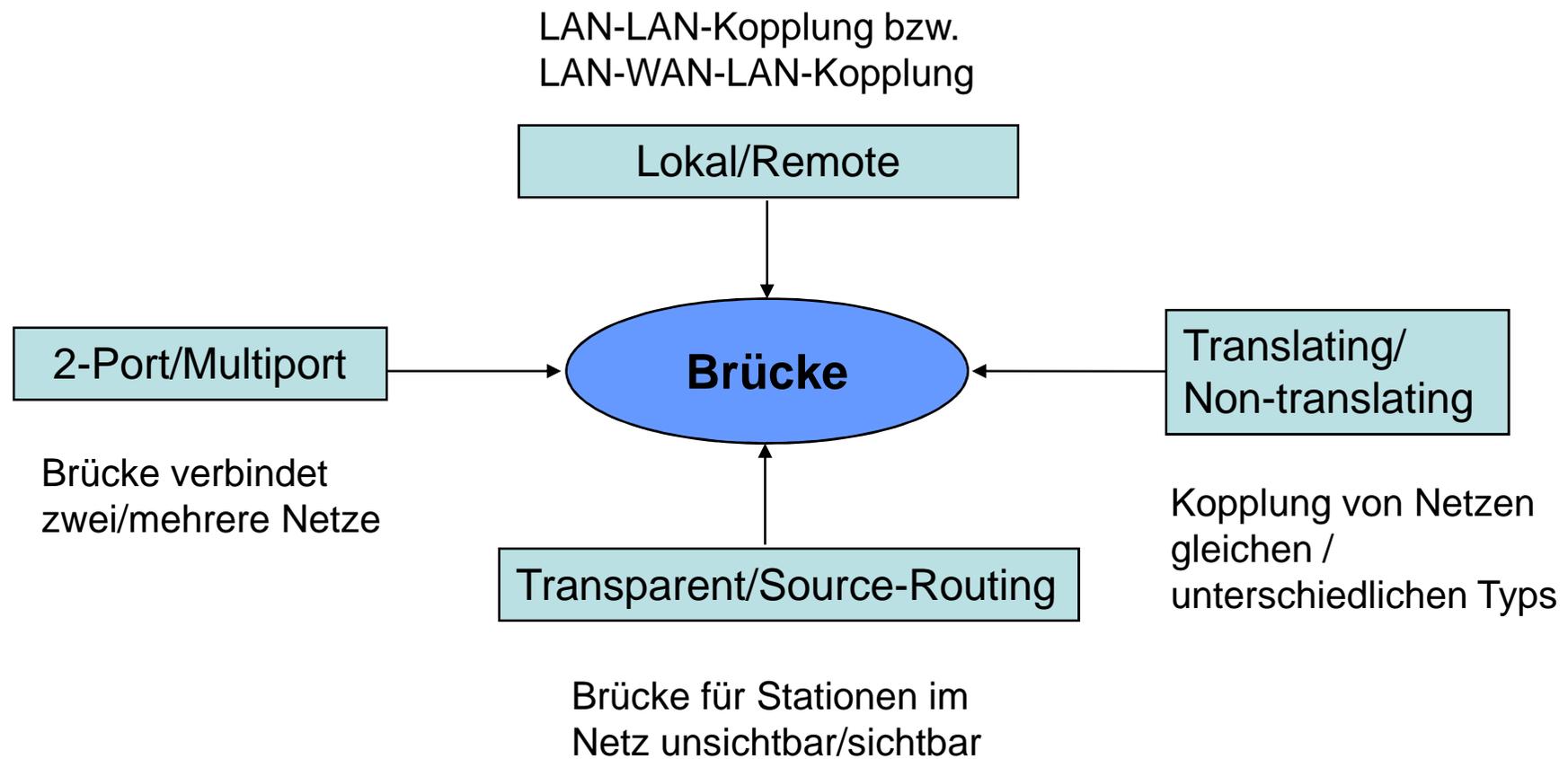
4.1.3. Brücke (Bridge)

- Kopplung von Netzen auf Schicht 2
 - Netzwerke vom gleichen Typ (z.B. 802.x mit 802.x) (non-translating)
 - Netzwerke unterschiedlichen Typs (z.B. 802.x mit 802.y ($x \neq y$)) (translating)
- Aufgaben:
 - Trennen des Intranetz-Verkehrs vom Internetz-Verkehr (Filterfunktion)
 - Erhöhung der Netzkapazität großer Netzwerke durch Partitionierung (jede Partition mit voller Stationszahl/Längenausdehnung)
 - Durchführung einfacher Wegewahlfunktionen (Einfache Vermittlung auch auf Schicht 2)





Brücken – Übersicht





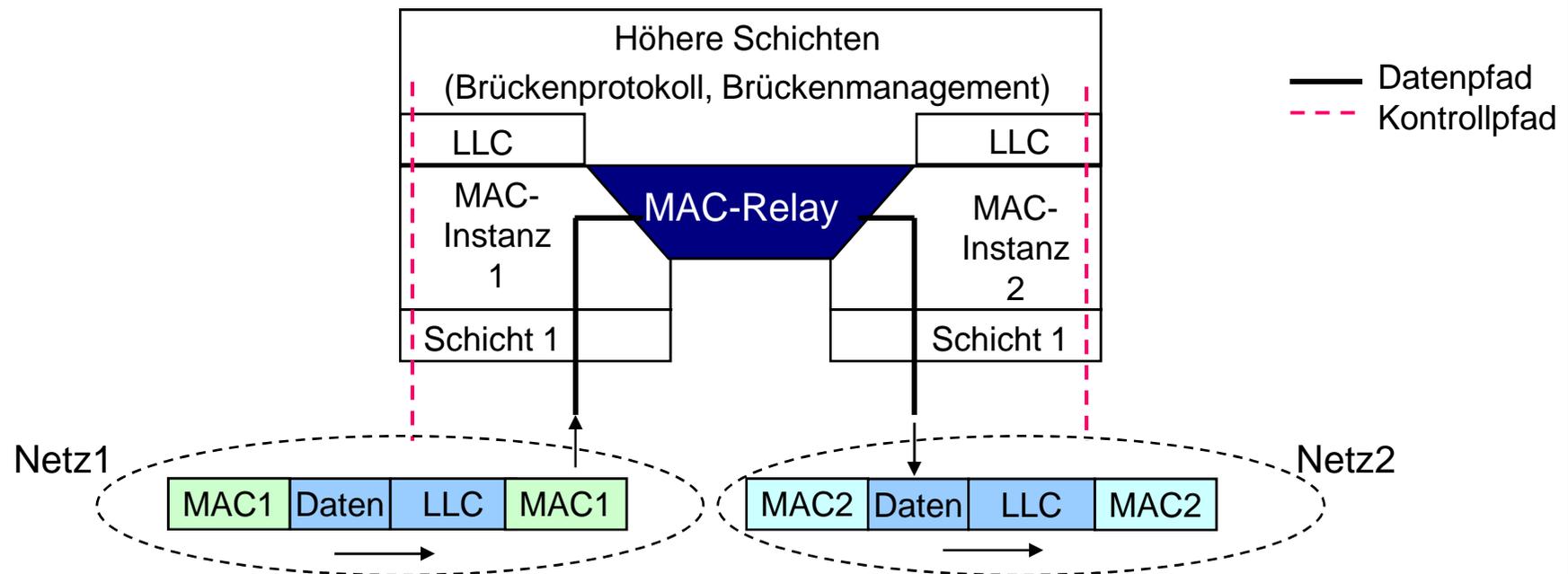
Typen von Brücken

- Zwei Basisvarianten von Brücken:
 - **Source-Route-Brücken (IEEE 802.5):**
 - Weiterleitungsinformation für die Brücke wird vom Endsystem in den Datenpaketen spezifiziert
 - Weiterleitung ist nicht transparent für das Endsystem
 - Endsystem muss Wegewahlinformation für die Dateneinheiten in Erfahrung bringen
 - Wenig Aufgaben innerhalb der Brücke, daher einfache Realisierbarkeit der Brücke
 - In der Praxis wenig eingesetzt
 - **Transparente Brücken (IEEE 802.1D):**
 - Weit verbreiteter Brückentyp
 - Weiterleitungsentscheidung wird von der Brücke eigenständig getroffen
 - Brücke verwaltet in der Regel eine Tabelle (die Filterdatenbasis)
 - Diese Information kann statisch angegeben werden
 - Typischerweise: dynamisches Lernen, indem die Brücke Information über die Lokation von Endsystemen sammelt, d.h. Adressen und zugehörige Ports lernt
 - Das Vorhandensein einer Brücke zum Zielsystem bleibt dem sendenden Endsystem verborgen
 - Endsystem ist nicht in die Wegewahl involviert



Transparente Brücke (MAC-Bridge)

- Merkmale:
 - Lokale, übersetzende Brücke
 - Für jedes Netzwerk eine eigene Schicht-1- und MAC-Instanz
 - Die MAC-Instanzen werden über ein MAC-Relay verbunden; dieses nimmt die Weiterleitungs- und Filterfunktion wahr
 - LLC-Instanzen nur für die höheren Schichten der Brücke (Brückenprotokoll, Brückenmanagement)

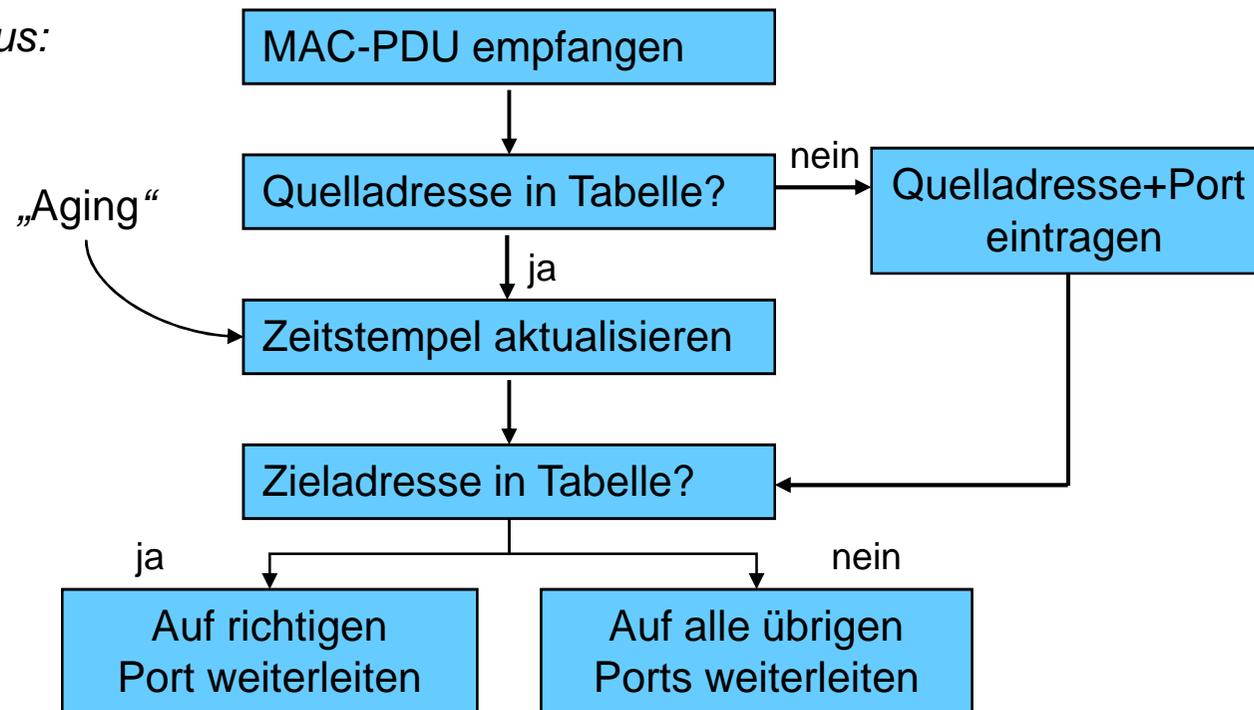




Transparente Brücke – Arbeitsweise

- Alle PDUs werden empfangen („promiscuous mode“)
- Brücke verwaltet eine Tabelle (forwarding database), in der sie Informationen über die Ausgangsleitung (Port) von Endsystemen sammelt (d.h. sie lernt MAC-Adressen)
- Das Vorhandensein einer Brücke zum Zielsystem bleibt dem sendenden Endsystem verborgen

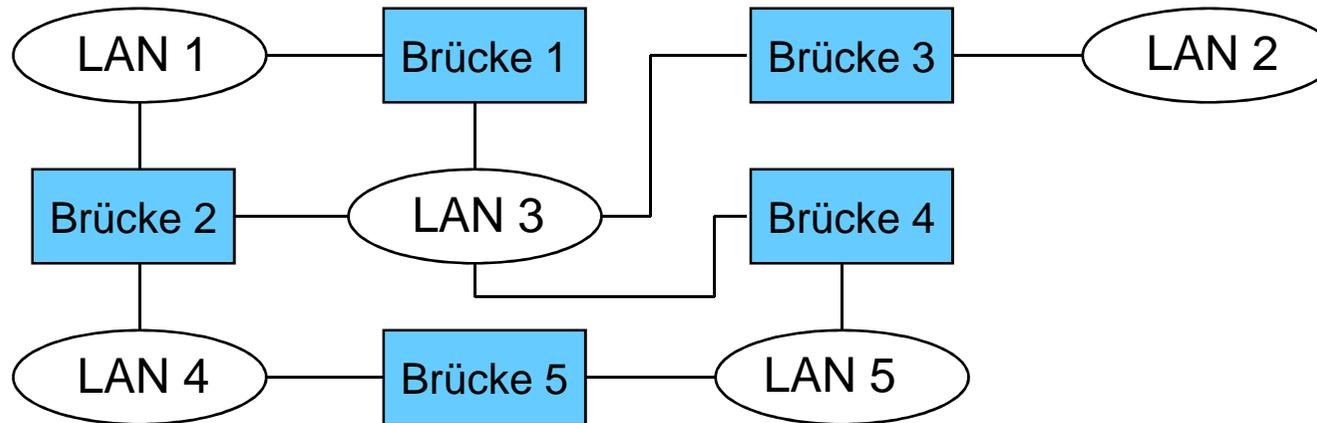
Forwarding-Algorithmus:





Brücken: Redundante Wege

- Bei brückengekoppelten Netzwerken können redundante Wege zwischen zwei Netzen existieren (z.B. zur Fehlertoleranz)



- **Probleme:**
 - Repliziert empfangene Datenpakete (über verschiedene Wege)
 - Endlos kreisende Datenpakete (Schleifen)
- **Lösung:**
 - Etablierung einer *logischen Baumstruktur* über allen Brücken der involvierten Netzwerke (⇒ Spanning-Tree-Algorithmus)
 - Weiterleiten von Datenpaketen nur entlang der Baumstruktur (eindeutiger Pfad), restliche Brücken blockieren ihre Ports

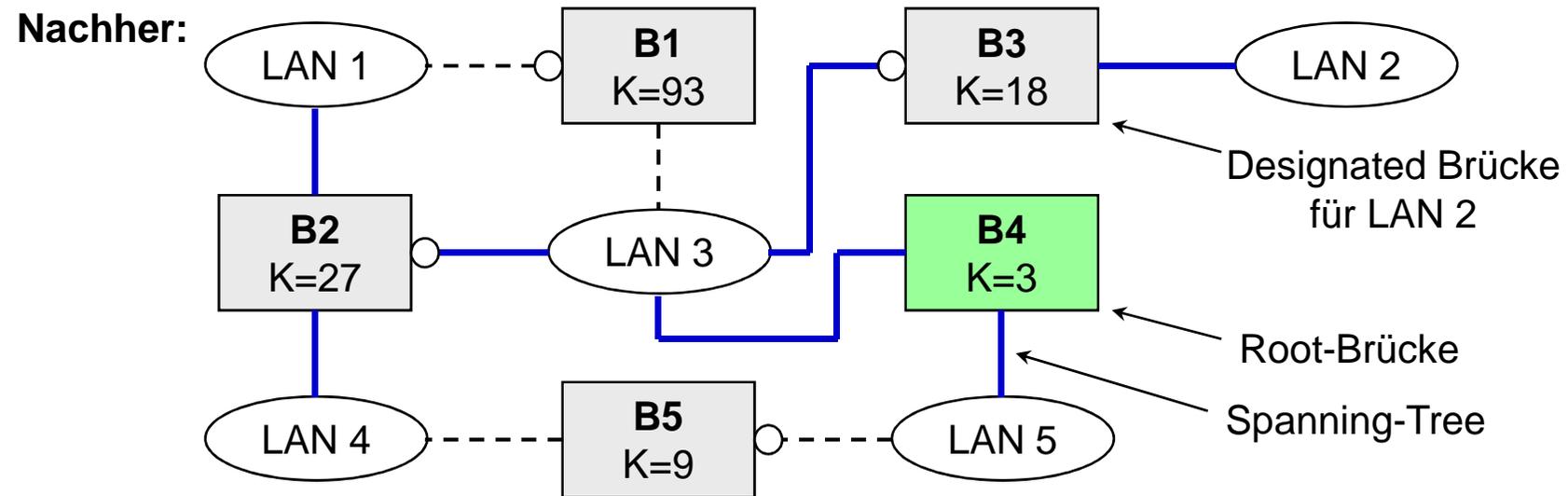
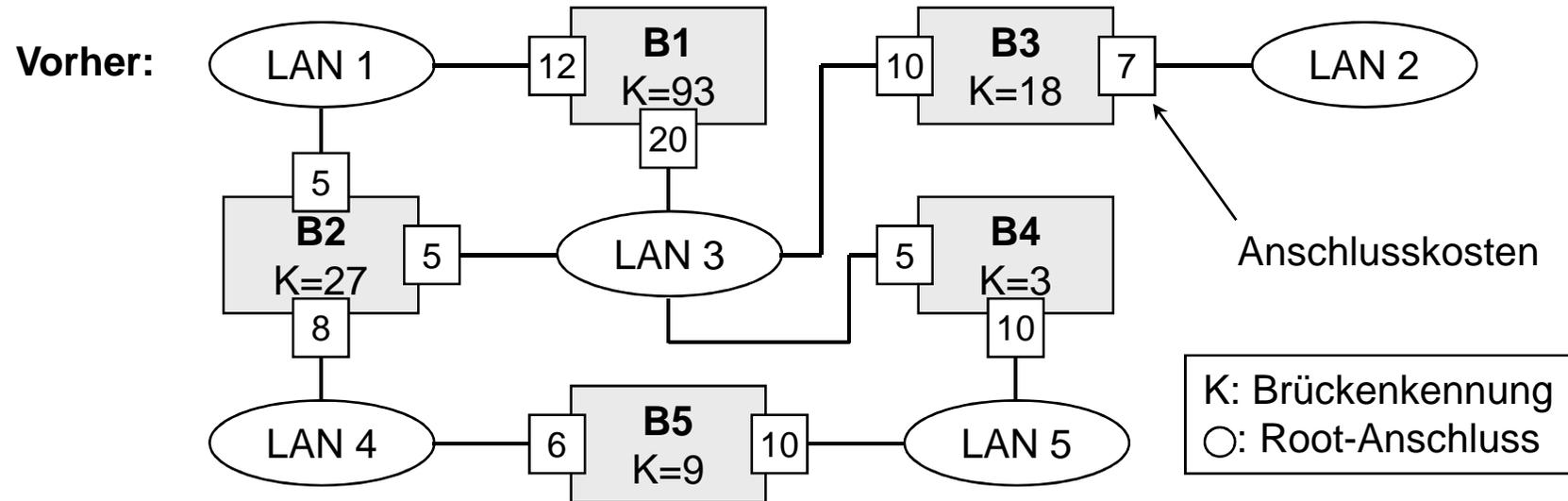


4.1.4. Spanning-Tree-Algorithmus

- Voraussetzungen:
 - Gruppenadresse zur Adressierung aller Brücken im Netzverbund
 - Eindeutige Brückenkennungen (MAC-Adresse)
 - Eindeutige Anschlusskennungen in jeder Brücke (MAC-Adresse)
 - Kosten an allen Anschlüssen einer Brücke („Anschlusskosten“)
- Ablauf:
 1. Bestimmen der Root-Brücke (Wurzel des Baumes):
 - Zuerst nimmt jede Brücke an, dass sie Root-Brücke ist
 - Root-Brücken senden regelmäßig Hello-Pakete mit ihrer Brückenkennung aus
 - Bei Erhalt eines Hello-Pakets mit kleinerer Brückenkennung ordnet sich eine Root-Brücke der anderen unter und sendet das Paket als Broadcast
 2. Bestimmen der Root-Ports
 - Root-Anschluss (Root-Port) einer Brücke: dies ist der Port, über den der günstigste Pfad Richtung Root-Brücke verläuft (nur Kosten für Ausgangsports werden berücksichtigt)
 - Summe über alle Anschlusskosten auf dem Weg zur Root-Brücke ist zu minimieren
 - Übertragungsgeschwindigkeit kann als Kostenfunktion dienen
 3. Bestimmen der Designated-Brücke:
 - Brücke mit günstigstem Root-Anschluss in einem Netzwerk wird als Designated-Brücke bestimmt
 - Root-Brücke ist Designated-Brücke für alle an sie angeschlossenen Netze



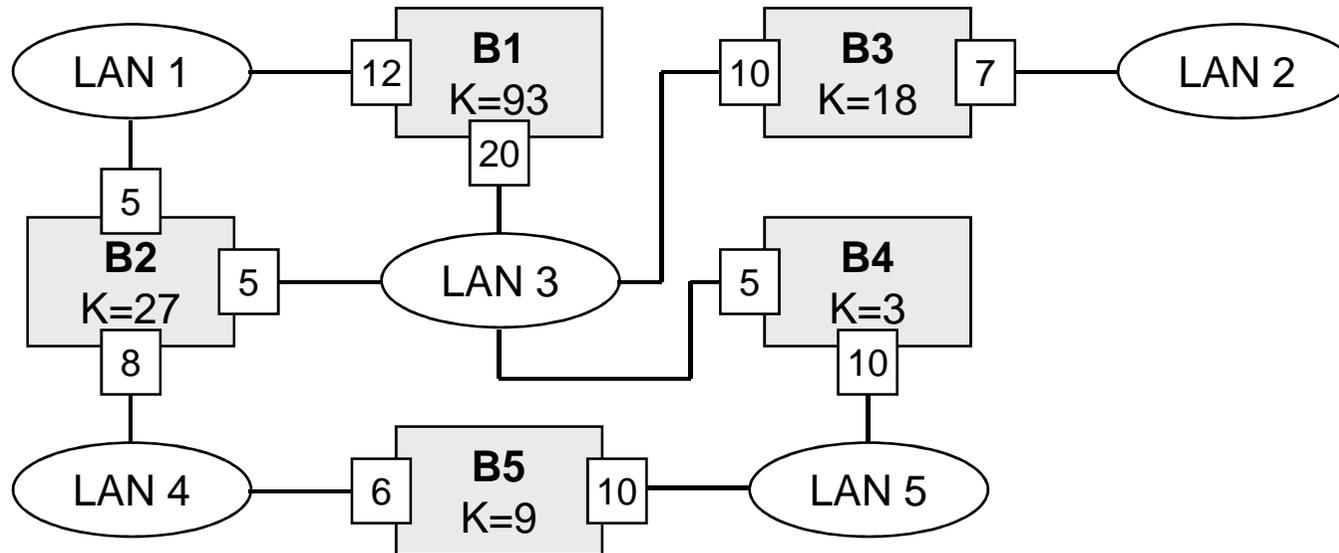
Spanning-Tree-Algorithmus: Beispiel





Spanning-Tree-Alg.: Berechnung der Root-Pfade

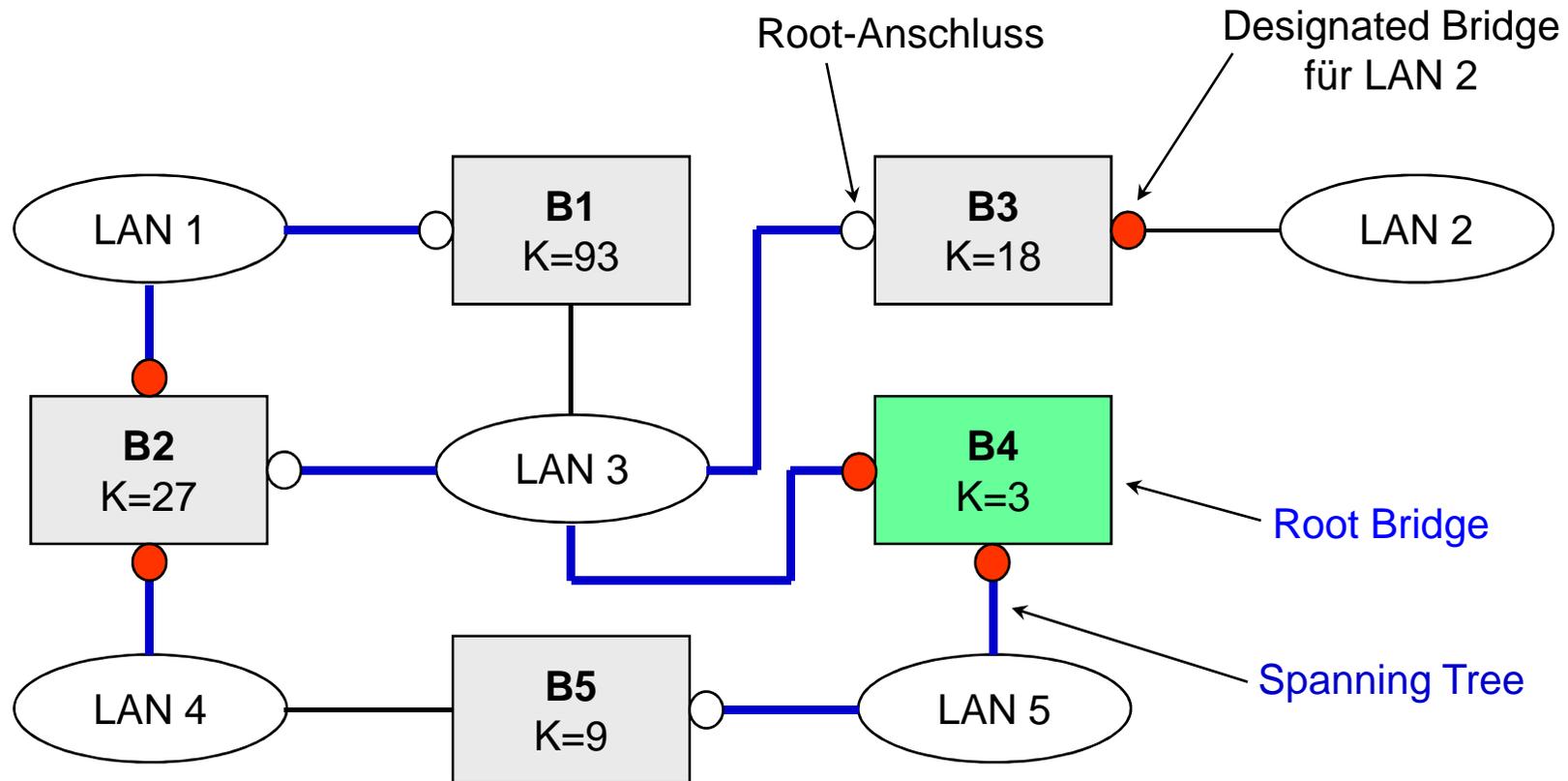
Vorher:



| Brücke | Kosten des Weges zur Root Bridge |
|--------|--|
| B3 | 10 (via LAN 3) |
| B1 | 20 (via LAN 3) 17 = 12 + 5 (via LAN 1 & LAN 3) |
| B2 | 5 (via LAN 3) 18 = 8 + 10 (via LAN 4 & LAN 5) 25 = 5 + 20 (via LAN 1 & LAN 3) |
| B5 | 10 (via LAN 5) 11 = 6 + 5 (via LAN 4 und LAN 3) |

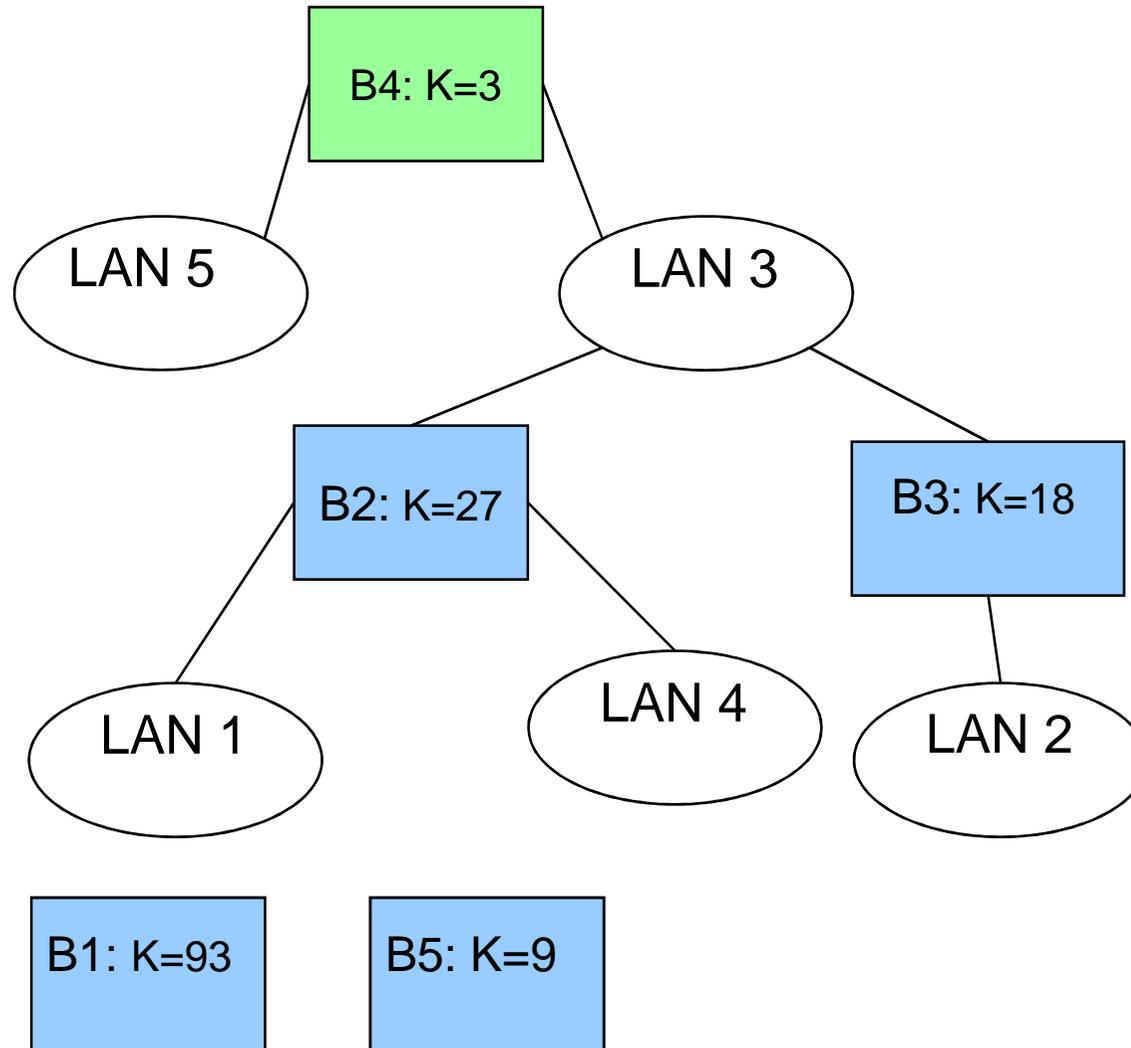


Spanning-Tree-Algorithmus: Bestimmung der Designated Bridges





Spanning-Tree-Algorithmus (Ergebnis)



| LAN | Designated Bridge |
|-------|-------------------|
| LAN 1 | B2 |
| LAN4 | B2 |
| LAN2 | B3 |
| LAN3 | B4 (Root Bridge) |
| LAN5 | B4 (Root Bridge) |
| --- | B5 |
| --- | B1 |



4.1.5. Remote-Brücken

- Kopplung von entfernten LANs über ein oder mehrere WANs

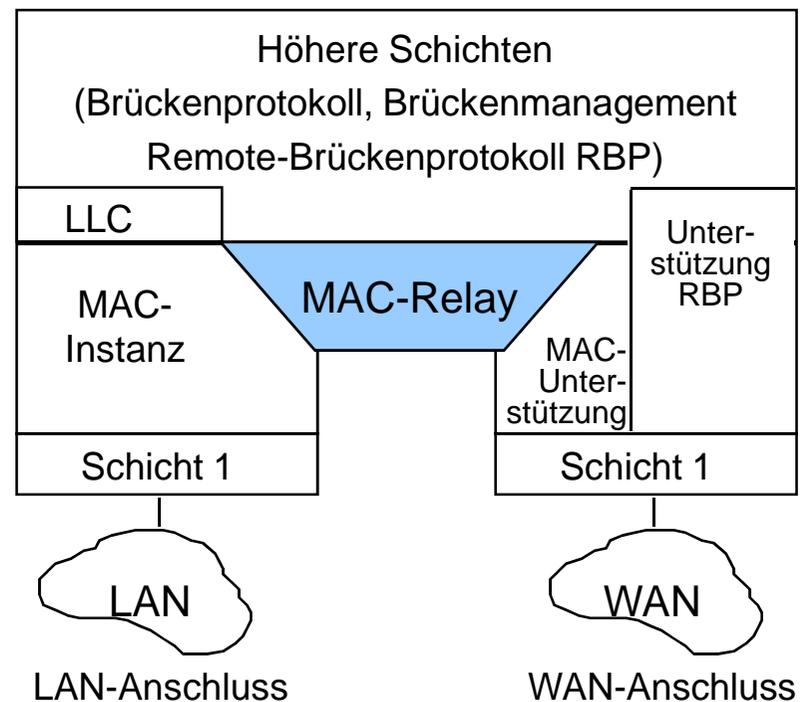


- Merkmale:
 - Einkapselung von MAC-Dateneinheiten (keine Umsetzung)
 - Remote-Brücken müssen paarweise auftreten
 - Nur zur Kommunikation zwischen LAN 1 und LAN 2 (Transparent)
 - Keine Kommunikation von LAN 1 (oder 2) mit dem WAN
 - Netzanschlüsse, die nicht mit einem LAN verbunden sind, werden als **virtuelle Anschlüsse** bezeichnet
- Beispiele:
 - Verbindung zweier Ethernets über ISDN
 - Verbindung zweier FDDI-Netze über ATM



Architektur einer Remote-Brücke

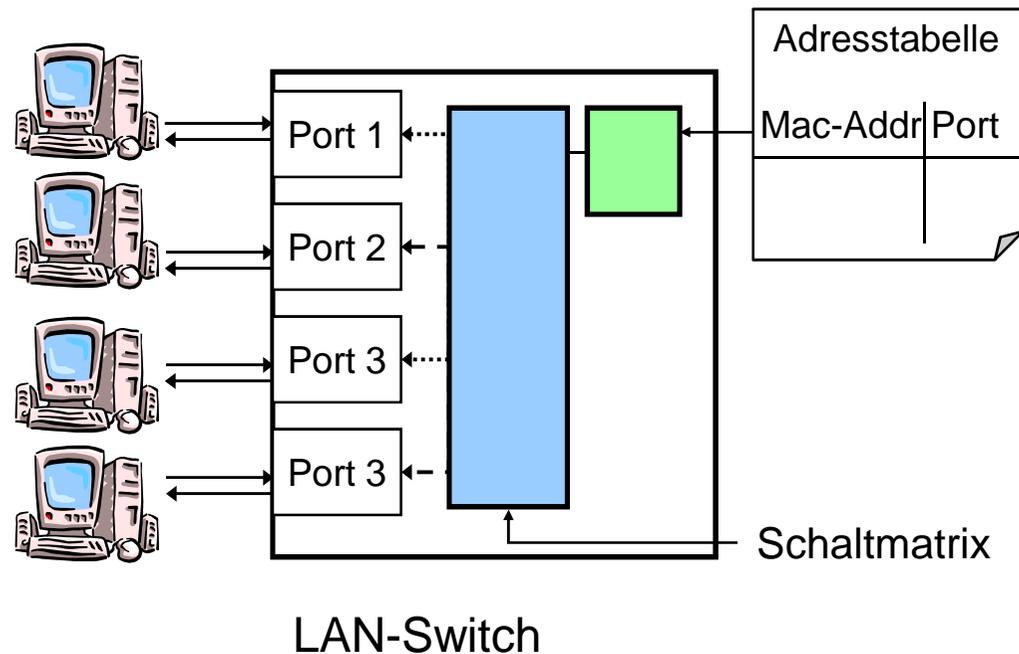
- Merkmale:
 - MAC-Unterstützung am virtuellen Anschluss für das MAC-Relay (zum WAN)
 - Remote-Brückenprotokoll (RBP) zwischen virtuellen Anschlüssen regelt die Punkt-zu-Punkt-Kommunikation der entfernten Brücken (über das WAN)
 - Brückenprotokoll setzt im WAN auf dem RBP auf





4.1.6. Switched LANs

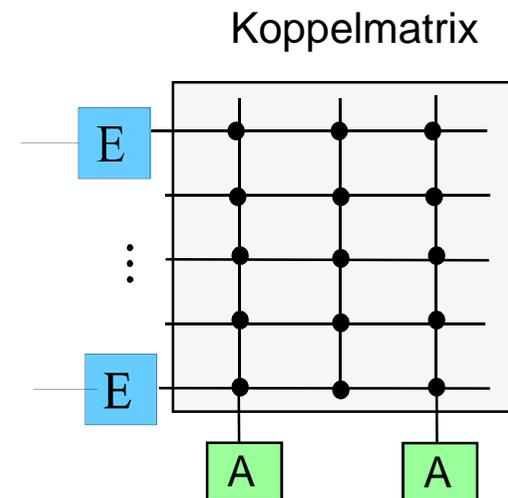
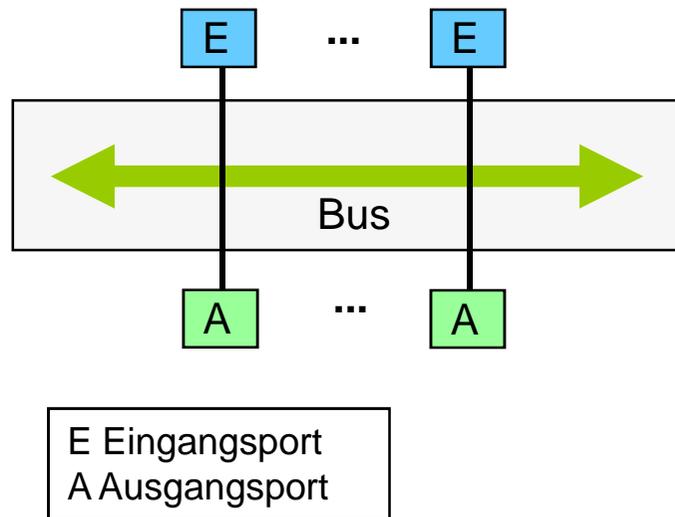
- Sternförmige Topologie, zentrale Komponente = *Switch*
- Jedes angeschlossene Gerät erhält die **volle Bandbreite**
- Sende-/Empfangsrichtung meist getrennt (Vollduplex-Übertragung)
- Parallele Bearbeitung/Weiterleitung mehrerer eingehender Rahmen
- Verbindung Eingang- zu Ausgangsport über Schaltmatrix anhand Adresstabelle





Schaltmatrix (switching fabric)

- **Aufgabe:** Schalten einer Verbindung zwischen 2 Ports (je Datenpaket)
- Realisierung in Hardware durch Application Specific Integrated Circuits (ASICs)
- Weiterleitung wird realisiert durch
 - ein-/mehrstufige Koppelmatrix, Bus, Ring, gemeinsamer Speicher
- Speicherelemente zur Konfliktauflösung notwendig
 - Eingangsspeicher, Ausgangsspeicher, verteilter Speicher



HoL - Head of Line Blocking



Switched LANs – Mechanismen

- Cut Through
 - Adresstabelle wird angesprochen, sobald die Zieladresse eingelesen ist
 - Weiterleitung des Datenpakets, sobald der Weg geschaltet ist
 - Geringe Latenzzeit

- Store and Forward
 - Datenpaket wird zunächst vollständig eingelesen und zwischengespeichert
 - Kontrolle der CRC-Prüfsumme und Ausführen von Filterfunktionen

- Hybrides Switching
 - Kombination von Cut Through / Store and Forward
 - Auswahl abhängig von Fehlerrate

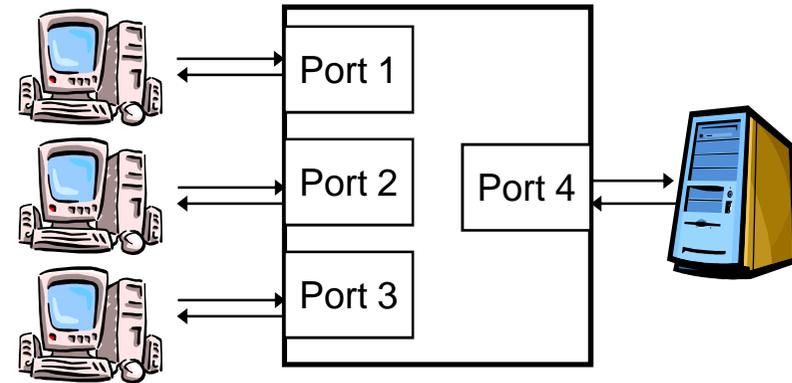
- Predictive Switching
 - Pfad in Schaltmatrix wird hergestellt, bevor Zieladresse vollständig eingelesen
 - Basierend auf den vorher geschalteten Pfaden



Ausprägungen

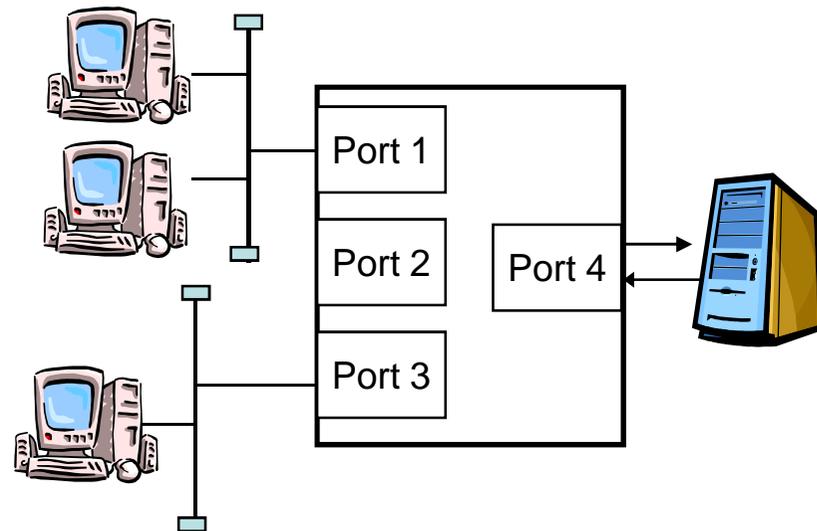
□ Port-Switching

- Nur ein Gerät pro Port (eine MAC-Adresse)
- Jedes Gerät erhält volle Bandbreite
- Schneller Table-Lookup möglich



□ Segment-Switching

- Mehrere Adressen pro Port zulässig
- Bandbreite wird jedem Segment zur Verfügung gestellt



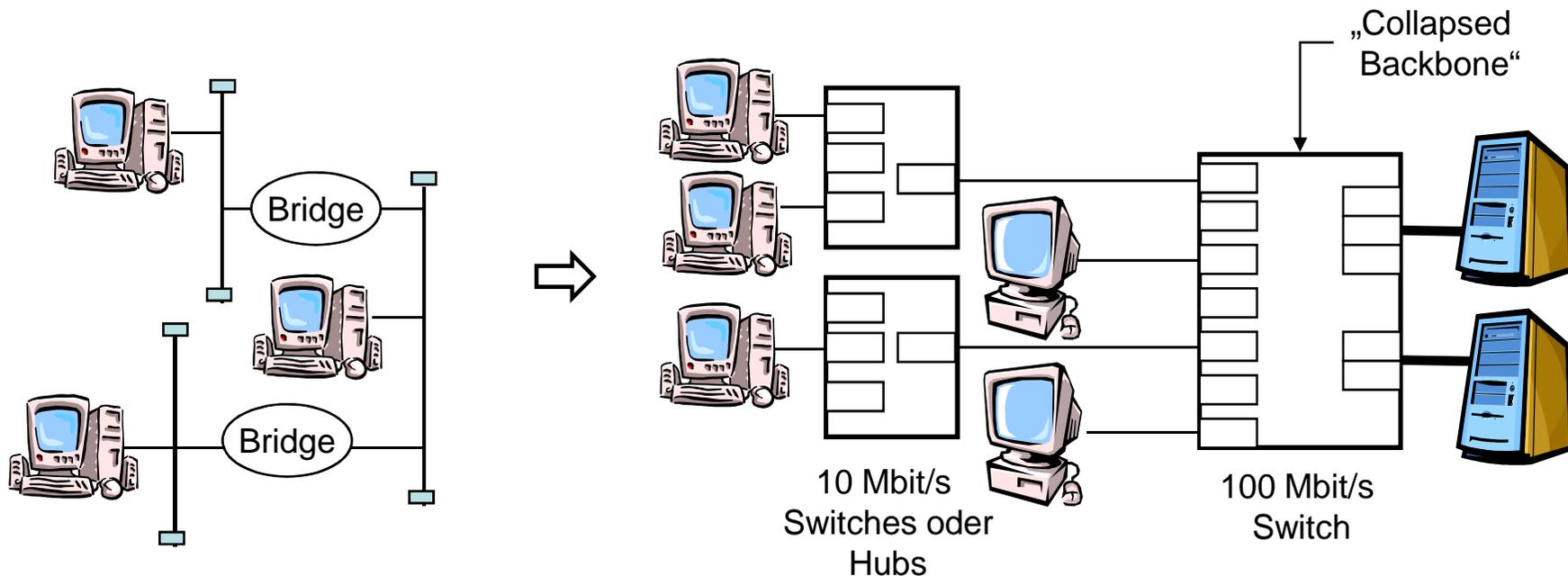
□ Bank-Switching

- Mehrere Ports teilen sich eine bestimmte Übertragungsrate



Einsatzmöglichkeiten

- Erforderliche Umstrukturierung des Netzwerks
 - Kopplung mehrerer LANs durch ein Hochgeschwindigkeits-LAN (Backbone)
 - Distributed Backbone ↔ Collapsed Backbone
 - Hochleistungs-Workstations an dediziertem Port
 - Server an mehreren Ports (direkter Anschluss an unterschiedlichen LANs, oder auch Port-Bündelung ⇒ Fat Pipe)





Weitere Switching-Verfahren

- Layer-3-Switching
 - Integration von Routing und Switching
 - Weiterleitung von Paketen anhand der Analyse der Felder des Schicht-3-Protokolls

- Layer-4-Switching
 - Erweiterung des Layer-3-Switchings um Analyse der Felder der Schicht 4 (z.B. Portnummern von TCP)
 - Priorisierung bestimmter Anwendungen

- IP Switching
 - IP-Switch = Kombination aus IP-Router (RIP, OSPF, BGP) und Layer-2-Switch
 - Zunächst konventionelles IP-Routing
 - Längeranhaltende Verkehrsflüsse werden erkannt, klassifiziert und gekennzeichnet
 - Unter Umgehung der Routing-Funktionalität wird danach auf Schicht 2 durch das Netz „geswitcht“, z.B. mithilfe von ATM
 - Heute wichtige Variante: MPLS – Multi-Protocol Label Switching

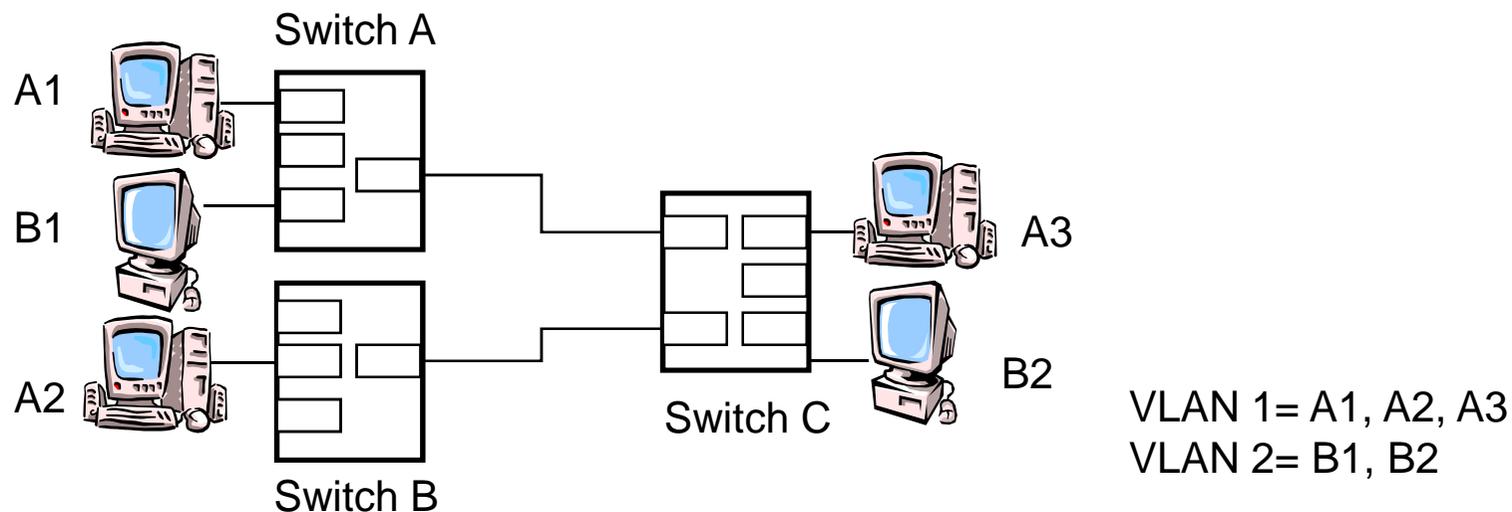


4.1.7. Virtuelle LANs (I)

□ VLAN

„Eine nach bestimmten Kriterien definierbare Broadcast-Domäne“

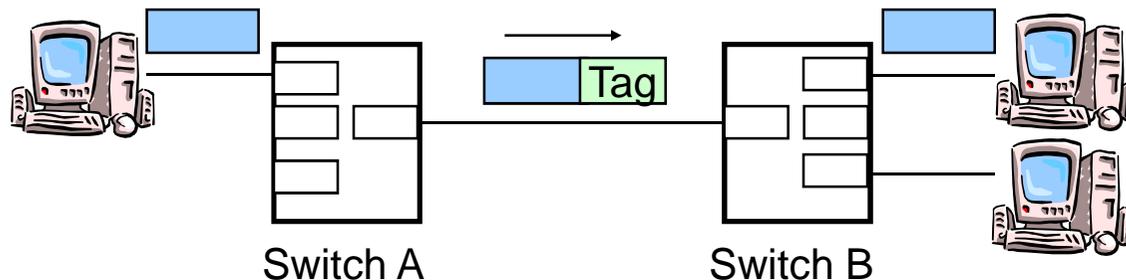
- Ziel: Trennung von physikalischer und logischer Netzwerkstruktur
 - Datenpakete werden ausschließlich innerhalb des jeweiligen VLANs verteilt
 - Mitglieder eines VLANs können räumlich verteilt sein, z.B. an verschiedenen LAN-Switches
 - ⇒ Unabhängigkeit von Standort und VLAN-Zugehörigkeit





Virtuelle LANs: Vorteile, Realisierung (I)

- Vorteile (allgemein)
 - Einschränkung von Broadcasts/Multicasts ⇒ bessere Ausnutzung der Bandbreite
 - Effizientere Verwaltung durch vereinfachte Konfiguration
 - z.B. bei Änderungen der Netztopologie durch Umzug
 - Erhöhte Sicherheit
 - Authentifizierung vor dem Beitritt einer Station zu einem VLAN
 - Strikte Trennung des Datenverkehrs verschiedener LANs
- Realisierung mit Tags
 - Analyse des eingehenden Pakets auf VLAN-Zugehörigkeit (interne Tabelle)
 - Erster Switch fügt ein „Tag“ an das Paket an (Kennung für jedes VLAN)
 - Erweitertes Rahmenformat – IEEE 802.1q: Tag aus vier Feldern, Länge von 32 Bit, wird nach MAC-Adress-Feldern eingefügt. Protokoll ID: zwei Byte (Wert: 0x8100); Prioritätenfeld: drei Bit, Indikator des Canonical Formats (signalisiert Darstellungsformat der Adressfelder): ein Bit; VLAN-ID: zwölf Bit.
 - Weiterleitung des Datenpakets an den nächsten Switch
 - Letzter Switch entfernt das Tag und übergibt das Paket an das Endsystem





VLAN: Realisierung (II)

□ *Schicht-2-VLANs*

- Realisierung durch LAN-Switches
- VLAN wird durch mehrere Ports festgelegt (*port-based VLAN*)
 - Pro Port können nur Stationen eines einzelnen VLAN angeschlossen sein
 - Mitgliedschaft in mehreren VLANs erfordert mehrere Netzwerkadapter
- VLAN durch eine Liste von MAC-Adressen definiert (*MAC-based VLAN*)
 - einfacher Umzug einzelner Stationen möglich

→ **Aber:** Router zur Kommunikation zwischen VLANs notwendig

□ *Schicht-3-VLANs*

- Realisierung durch Layer-3-Switches
 - integrierte (aber nicht notwendigerweise vollständige) Schicht-3-Fähigkeit vorhanden!
- VLAN wird durch Subnetz-Adresse festgelegt (*subnet-based VLAN*)
- VLAN wird durch Netzwerkprotokoll festgelegt (*protocol-based VLAN*)

→ Kein zusätzlicher Router zur Kommunikation zwischen VLANs notwendig



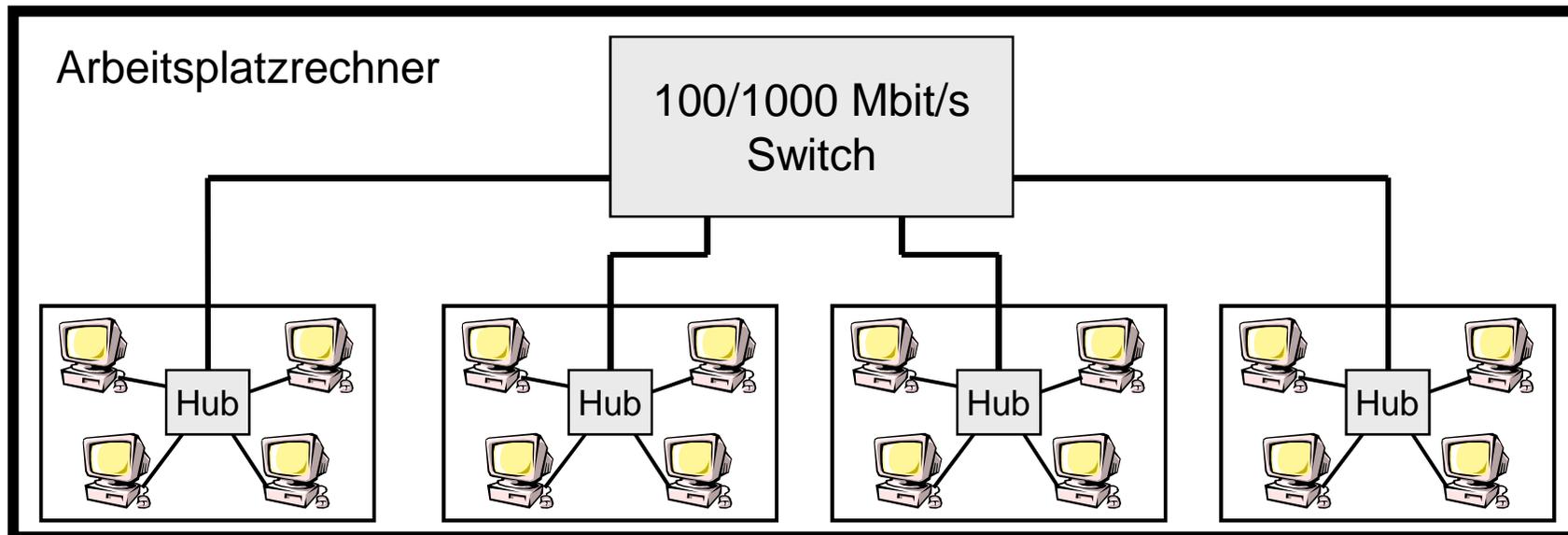
VLAN: Realisierung (III)

- *Regelbasierte VLANs*
 - Beliebige Verknüpfung von Feldern der Schichten 2 und 3 zur Definition eines VLAN
 - Beispiel für die Definition regelbasierter VLANs
 - VLAN 1 = „Alle IP-Benutzer innerhalb eines bestimmten Subnetzes“
 - VLAN 2 = „Jeglicher Datenverkehr mit einem bestimmten Wert des Typ-Feldes im Ethernet-Header“
 - VLAN 3 = „Alle Rechner, deren Netzwerkkarte vom gleichen Hersteller stammt“
 - Vorteil:
 - besonders flexible Konfigurationsmöglichkeit
 - Nachteile:
 - Aufwendige Einrichtung der VLANs
 - Erhöhte Latenzzeit durch die Abarbeitung der einzelnen Regeln



4.1.8. Leitbeispiel: Strukturierte Verkabelung

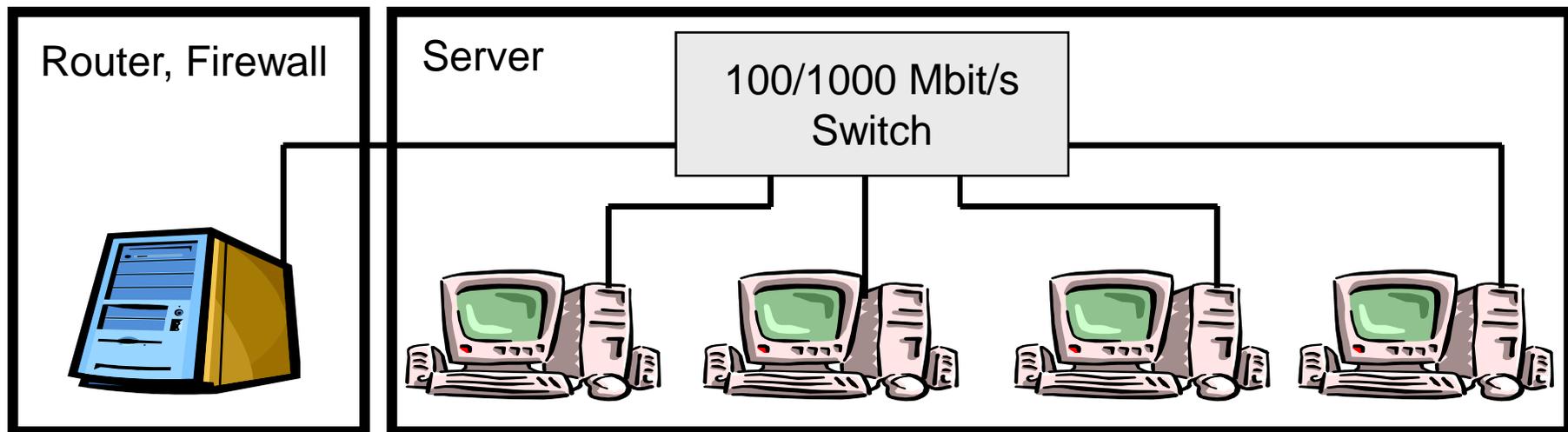
- Kopplung der einzelnen Zimmer eines Stockwerkes
 - Hub wenig leistungsfähig, da es zu häufigen Kollisionen kommen kann, wenn mehrere Rechner untereinander kommunizieren bzw. auf einen Server zugreifen
→ Switch
 - Switch arbeitet i.d.R. mit Fast-Ethernet, zum Teil mit Gigabit-Ethernet-Komponenten
– abhängig von den benutzten Anwendung.
 - Verkabelung mit Twisted-Pair-Kabel (UTP-5)





Leitbeispiel: Strukturierte Verkabelung

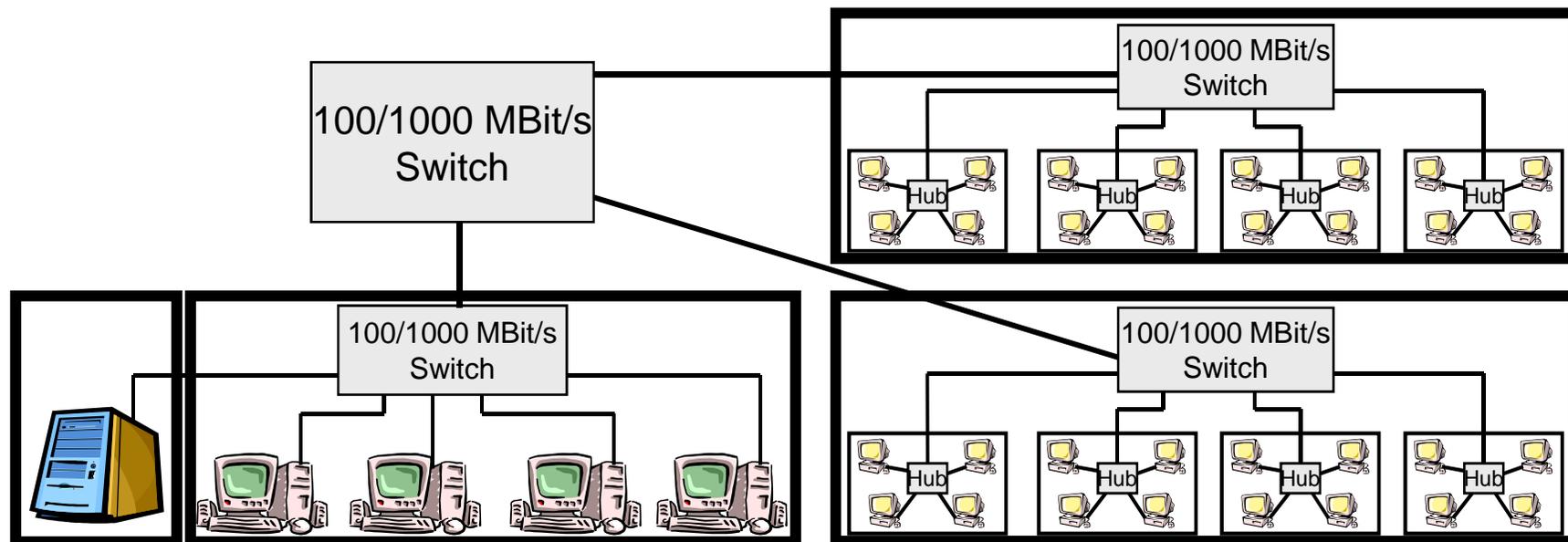
- Verkabelung der Server im Keller
 - Server werden i.d.R. stark frequentiert, daher kann ein Hub zu einem Engpass werden
 - Daher werden Server häufig direkt an Switches angeschlossen
 - Fast-Ethernet- bzw. Gigabit-Ethernet-Switch
 - Ggf. kann es sinnvoll sein, Glasfasern anstelle von UTP-5-Kabeln einzusetzen.
 - Alternativ kann statt des Routers einen Layer-3/4-Switch verwendet werden, um z.B. die über den Router kommende HTTP-Anfragen schnell an den richtigen Server zu leiten





Leitbeispiel: Strukturierte Verkabelung

- Verkabelung der Stockwerke
 - Kopplung einzelner Stockwerke über 100/1000 Mbit/s Switch
 - Alternative: Anschluss der Switches der (oberen) Stockwerke am Server-Switch (im Keller) sofern dort ausreichend Kapazitäten vorhanden sind
 - Einzelne Rechner können mittels VLANs zu einem log. Netzwerk gekoppelt werden





Inhalt des 2. Teils des Kapitels

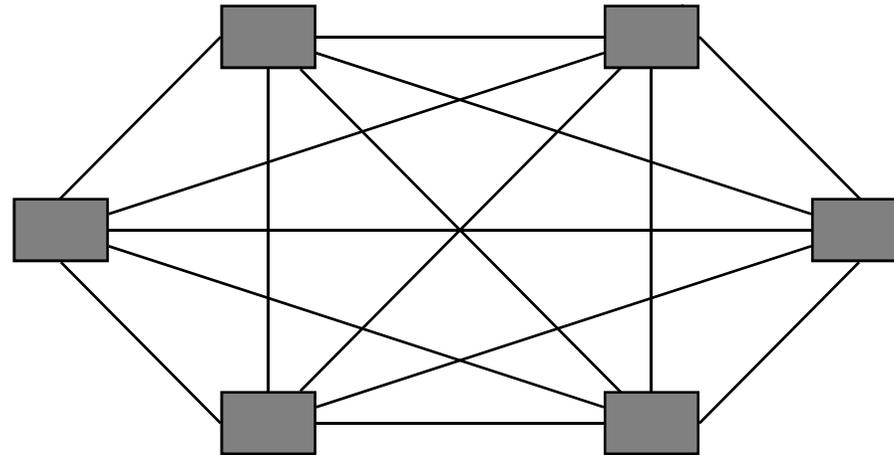
4.2. Vermittlungsprinzipien für globale Netze (Schicht 3)

- Leitungsvermittlung
- Nachrichten-/Speichervermittlung
- Paketvermittlung
 - virtuelle Verbindung
 - Datagrammvermittlung
- Router
- Routing-Verfahren (Wegewahl)



Grundanforderungen and Telekommunikationsnetze

- Teilnehmer sollen temporäre Kommunikationsbeziehungen mit anderen Teilnehmern auf Anforderung durchführen können.
- Ein für jeweils zwei Teilnehmer permanent vorgehaltener Kommunikationsweg führt zu einer nicht beherrschbaren Zahl von Kommunikationswegen.
- Bei voller Vermaschung von N-Teilnehmern: $N * (N-1)$ Verbindungen



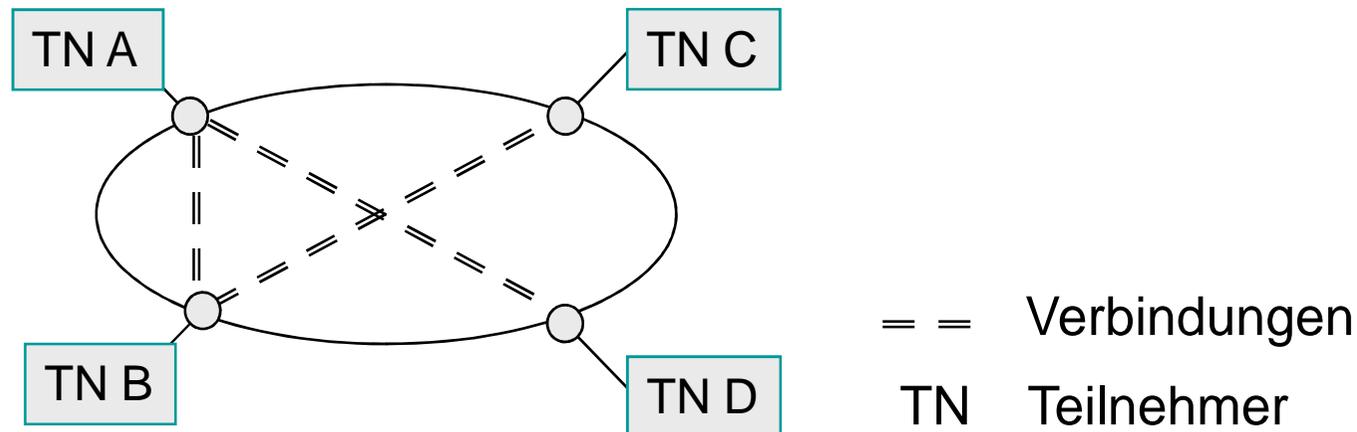
Vollständig vermaschtes Netz aus vielen überlagerten Punkt-zu-Punkt-Verbindungen

- In großen Netze (großen Anzahl Endsysteme, vielen Vermittlungssysteme) kommen aus Gründen mangelnder Leistungsfähigkeit die z.B. von Brücken bekannten Verfahren zum Einsatz, sondern spezielle Routing-Protokolle



Teilnehmersicht eines TK-Netzes

- Verbindungsorientiert - ein Teilnehmer unterhält zu jedem *Kommunikationspartner* eine Verbindung.
 - Hinweis: Der Begriff der Verbindung wird in der Telekommunikation in vielfältiger und sehr unterschiedlicher Weise genutzt.
 - Der spezielle Kontext definiert die notwendigen Details, beispielsweise die Schicht-3- oder physikalische Verbindung.
- Aufgaben, die mit einer Verbindung anfallen:
 - *Herstellen* von Verbindungen durch Vermittlungsdienste
 - Netzinterne *Überwachung* von *Verbindungseigenschaften* durch Monitore





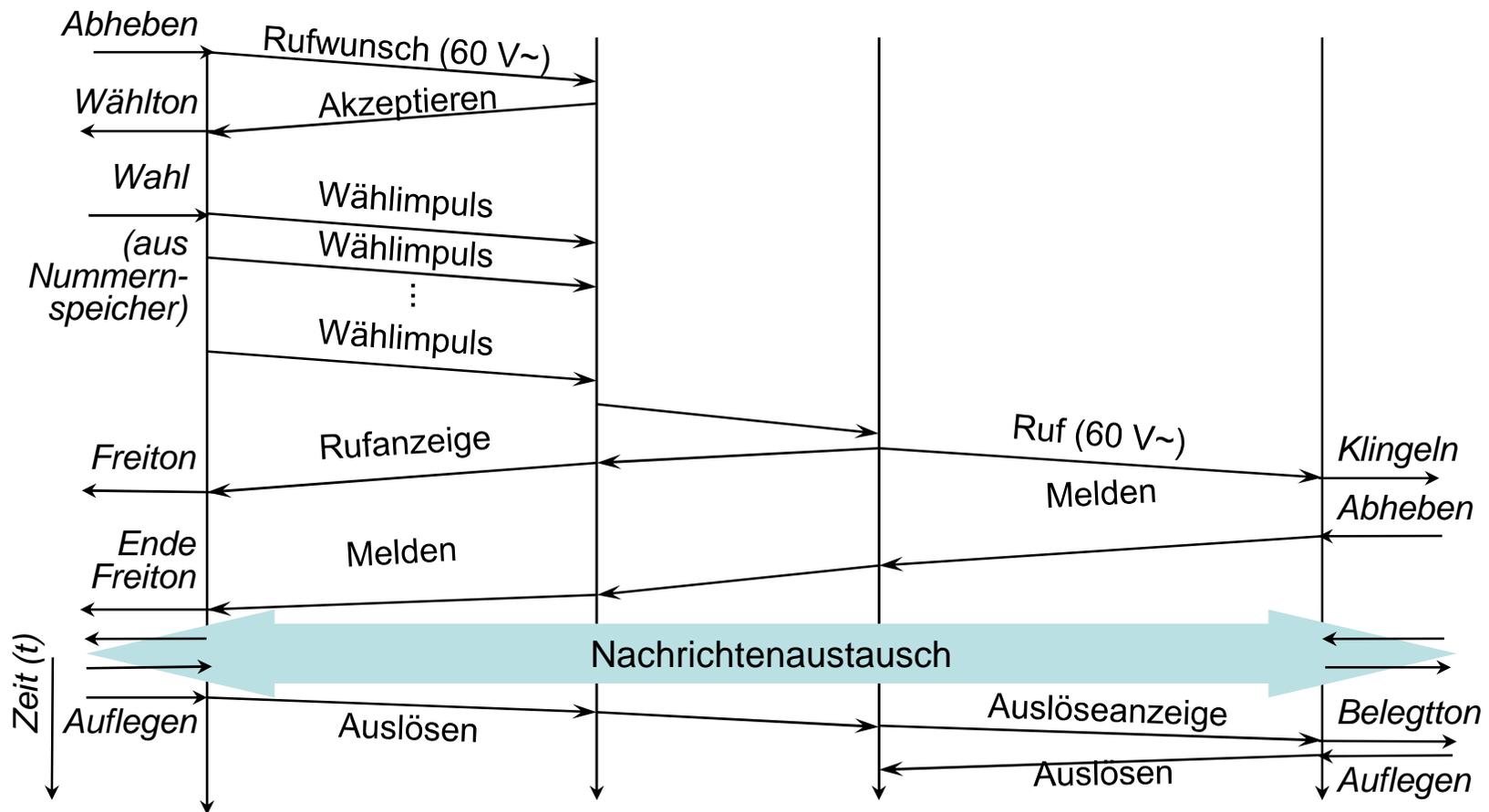
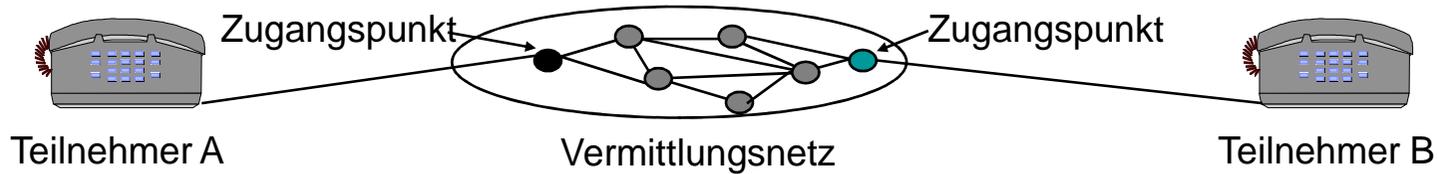
Vermittlung und Signalisierung

- Vermittlung
 - Grundlegende Aufgabe
 - Bereitstellung eines temporären Kommunikationsweges durch das Netz auf Anforderung
 - Geschieht über eine Kette von Vermittlungsstellen zum Zwecke des Datenaustausches zwischen zwei oder mehreren Teilnehmern
 - Arten
 - Verbindungsorientierte Vermittlungstechniken
 - Verbindungslose Vermittlungstechniken

- Signalisierung
 - Oberbegriff für die vermittlungstechnische Kommunikation
 - zwischen Teilnehmereinrichtung (Endgerät) und Netz
⇒ User-Network-Interface (UNI)
 - sowie im Netzinernen, also zwischen Netzknoten
⇒ Network-Network-Interface (NNI)



Wiederholung: Signalisierung im analogen Fernsprechnet





Verbindung

- Verbindung oder Netzwerkverbindung
 - Koppelt in verbindungsorientierter Art zwei Teilnehmer über eine oder mehrere Vermittlungsstellen (bei Wählnetzen temporär) zum Zwecke des Datenaustausches
 - Arten von Teilnehmern
 - Rufender Teilnehmer: der eine Schicht 3-Verbindung wünschende A-Teilnehmer (Anrufer)
 - Gerufener Teilnehmer: der eine Schicht 3-Verbindung erhaltende B-Teilnehmer (Angerufener)
 - Prinzip des Verbindungsaufbaues
 - Rufender Teilnehmer liefert neben dem Verbindungswunsch auch Identifizierungsdaten für den gerufenen Teilnehmer
 - Rufnummer, Kennung, Dienst-Nummer, Adresse, usw.
 - Zwischen Vermittlungsstellen werden zusätzlich vermittlungstechnische Daten ausgetauscht (Signalisierung)

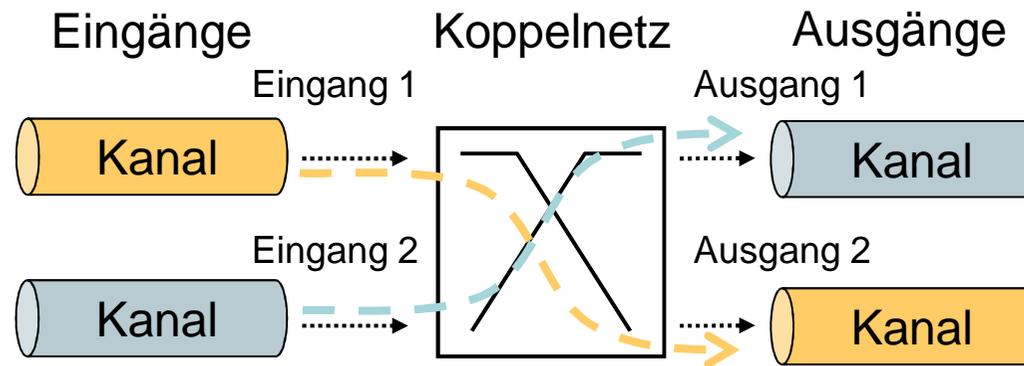


4.2.1. Vermittlungsprinzip Leitungsvermittlung

- Synonyme Begriffe
 - Leitungsvermittlung
 - Durchschaltvermittlung
 - Circuit Switching
 - Line Switching
- Lange Zeit dominierende Vermittlungstechnik für öffentliche Fernmeldenetze
- Aufbau eines durchgehenden, nicht-speichernden Übertragungskanal (Leitung) zwischen den Teilnehmern
- Übertragungsverzögerungen sind auf physikalisch bedingte signaltechnische Laufzeiten beschränkt
- Bitfolgen werden reihenfolgetreu übertragen (signalübertragungstechnisch bedingt), damit wird die Absendereihenfolge beim Empfänger beibehalten (wire-like feature)



Leitungsvermittlung (Circuit Switching)



- Koppeltabelle:

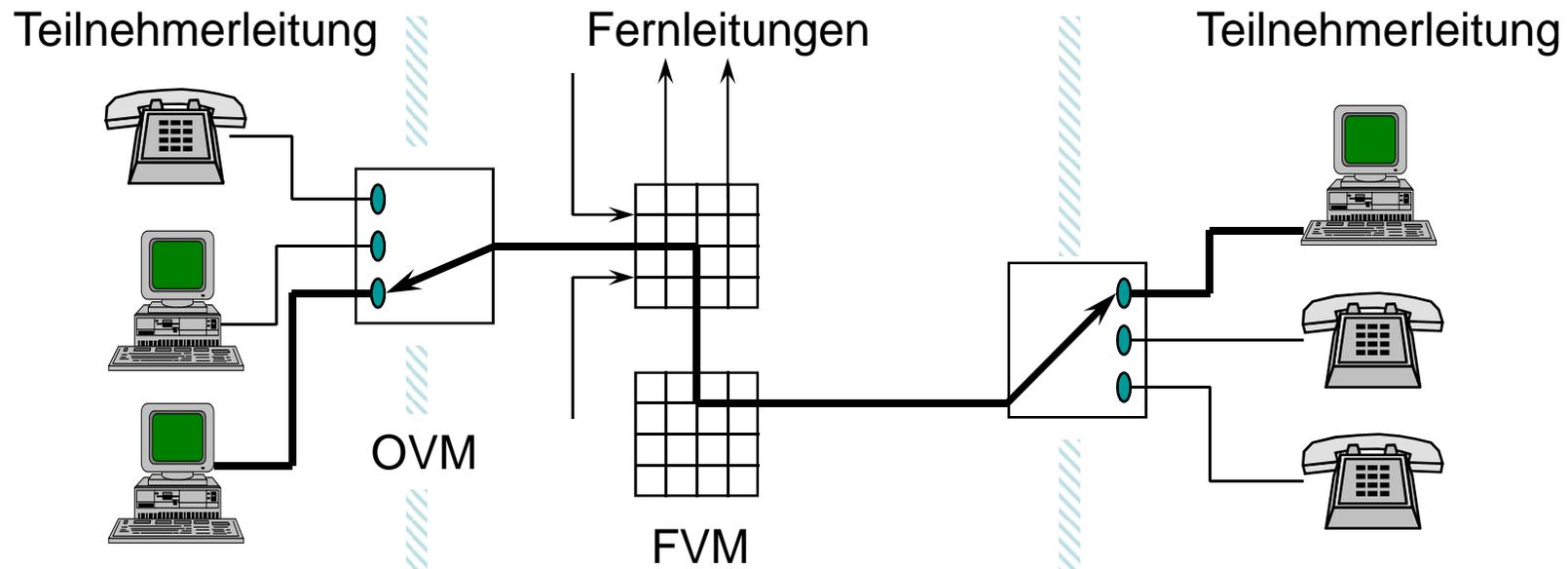
| Eingang | Ausgang |
|---------|---------|
| 1 | 2 |
| 2 | 1 |
| ... | ... |

- Kopplung physikalischer Eingangs- und Ausgangskanäle
 - mit Kanal ist nicht allein eine Leitung gemeint, sondern auch Zeitschlitz bei TDM, Träger bei FDM,...
- Verbindungsaufbau erforderlich → Einträge in der Koppeltabelle
- Beispiel: Leitungsvermittelltes Telefonnetz



Leitungsvermittlung im Telefonnetz

- Schema eines Durchschalttechnik-Vermittlungssystems

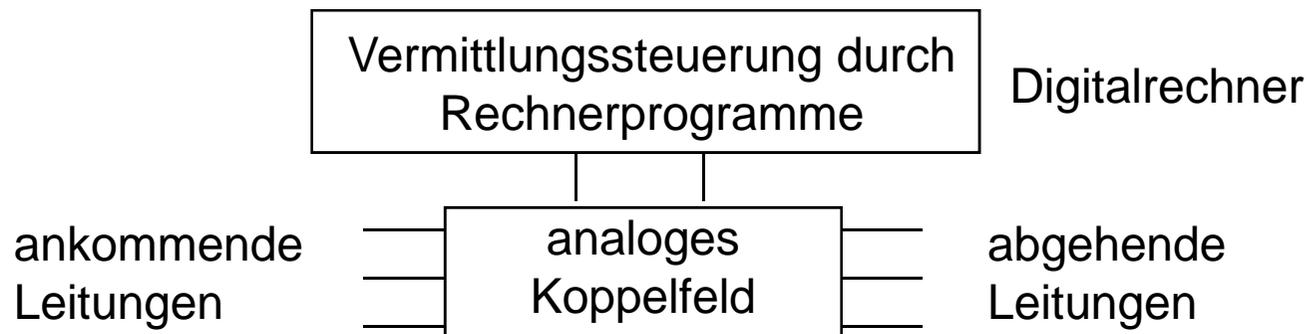


OVM Ortsvermittlung
FVM Fernvermittlung



Historie der Leitungsvermittlung

- Klassische elektromechanische Wähler (überholt)
 - Heb-Dreh-Wähler
 - Edelmetall-Motor-Drehwähler (EMD)
- Rechnergesteuerte Vermittlungen (Stored Program Control, SPC)
 - Beispiel: Rechnergesteuertes Vermittlungssystem mit analogem Koppelfeld



- Volldigitale Vermittlungssysteme
 - Digitalisierung der Vermittlungsstellen zusammen mit der digitalen Übertragungstechnik war Voraussetzung für das ISDN



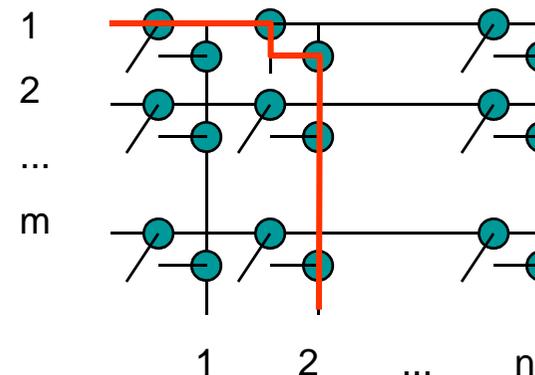
Leitungsvermittlung bei Raummultiplex

- Synonyme Begriffe
 - Raummultiplex
 - Raumvielfach
- Ein- und Ausgangskanäle sind physikalische Leitungen
- Durch Schließen eines Koppelpunktes wird die Verbindung hergestellt.
- Für analoge Übertragungen können ausschließlich Raummultiplex-Koppelanordnungen (Koppelfelder) verwendet werden.

- Beispiel

- m Zubringerleitungen
- n Abnehmerleitungen
- Hier:
Zubringer 1 mit
Abnehmer 2 verbunden

Zubringer

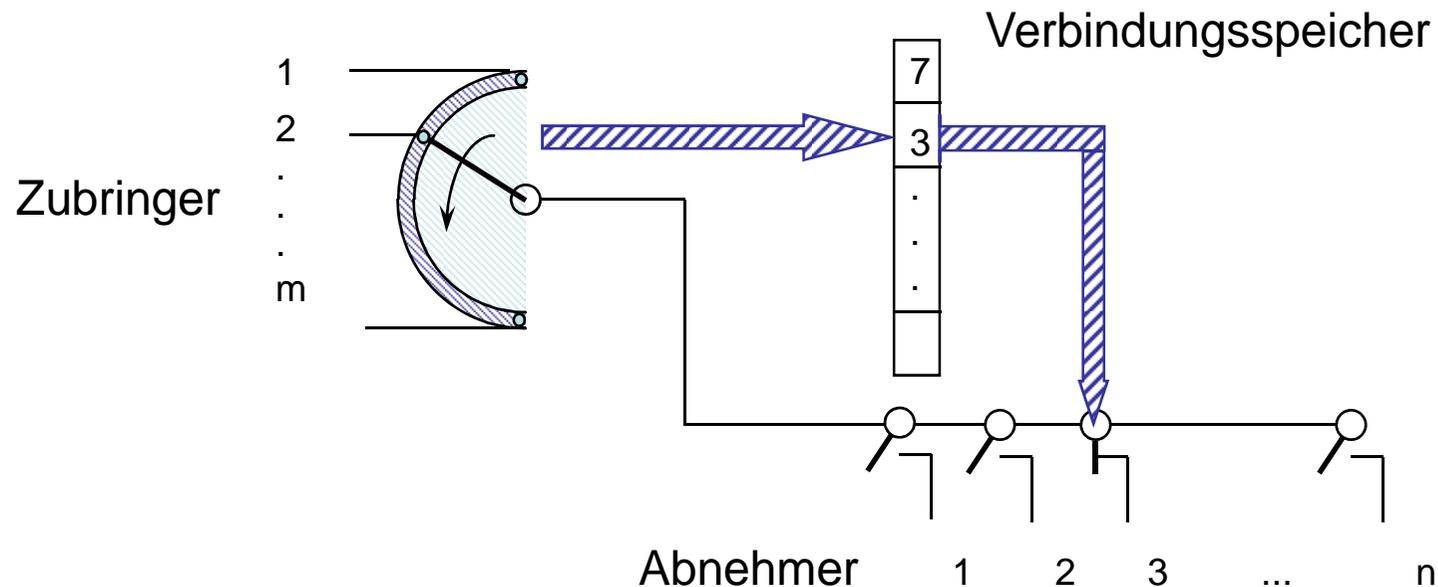


Abnehmer



Leitungsvermittlung bei Zeitmultiplex

- Synonyme Begriffe
 - Zeitmultiplex
 - Time Division Multiplex
- Kanäle entsprechen festen Zeitschlitzten im Zeitmultiplexschema
- Der Zubringer ist mit dem Abnehmer nur für kurze Abtast-Zeitintervalle über den Koppelbus verbunden.
- Der Verbindungsspeicher hält die Zuordnung „Zubringerleitung – Abnehmer“ für die jeweiligen Abtastzeitpunkte.



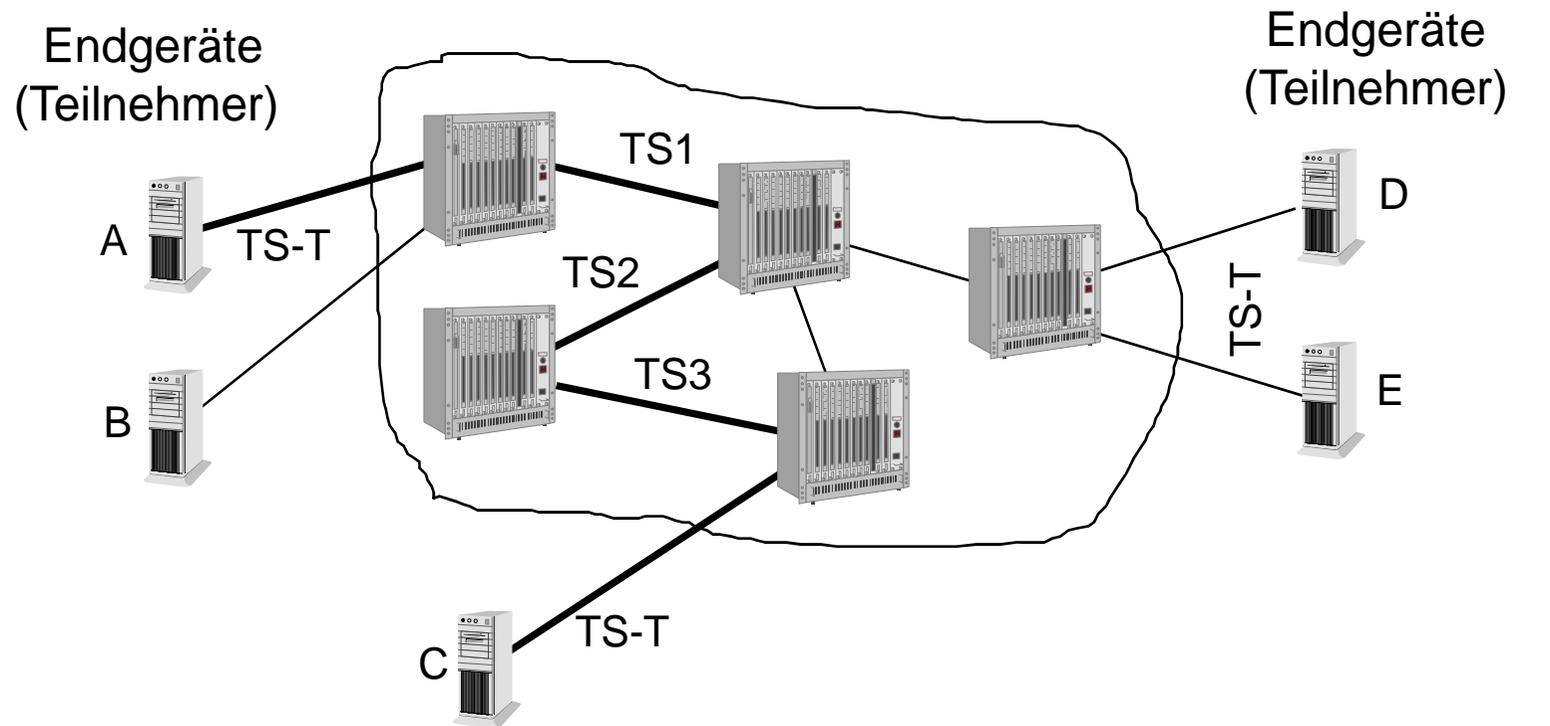


4.2.2. Vermittlungsprinzip Nachrichtenvermittlung

- Synonyme Begriffe
 - Nachrichtenvermittlung
 - Speichervermittlung
 - Teilstreckenvermittlung
- Älteste Vermittlungstechnik in Form der Telegrammvermittlung
- Eigenschaften:
 - Speicher (Puffer) in den Rechnern der End- und Durchgangsvermittlungsstellen
 - Es treten keine Besetzt- oder Belegtfälle, sondern nur Wartefälle auf.
 - Zwischengespeicherte Vermittlungsdatenblöcke warten auf das Freiwerden der gewünschten Teilstrecke.
 - Verlust von Vermittlungsdatenblöcken durch Speicherüberlauf möglich
 - Wartende Vermittlungsdatenblöcke werden in Speichervermittlungsstellen z.B. nach Prioritäten oder „zufällig“ umgeordnet vermittelt (Reihenfolgevertauschung)
 - Es besteht i. A. keine feste Zeitbeziehung zwischen den einzelnen Speichervermittlungsdatenblöcken.
 - Geschwindigkeitsanpassung zwischen unterschiedlich leistungsfähigen Endgeräten ist möglich



Speichervermittlungsnetz: Topologische Sicht



Vermittlungsrechner

TS-T

Teilstrecke-Teilnehmeranschluss



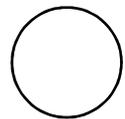
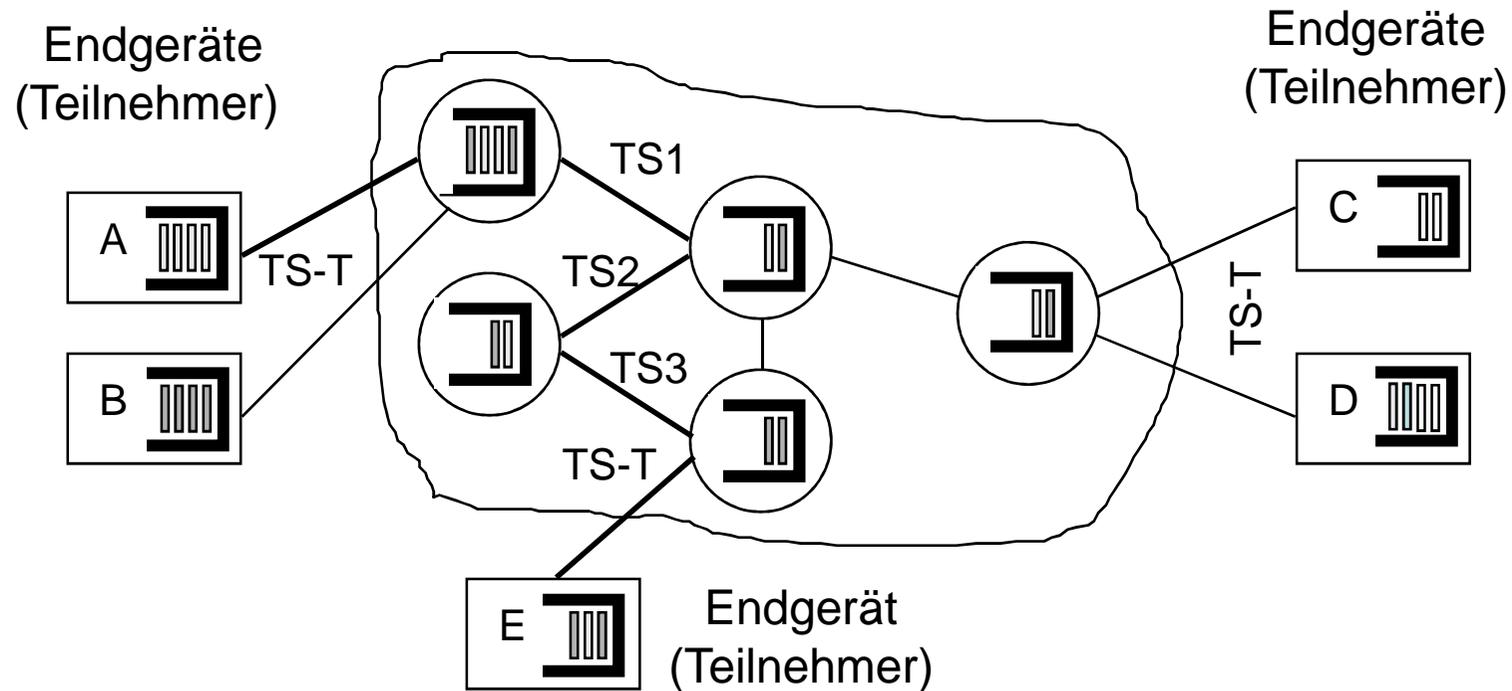
Endgerät (Teilnehmer)

TSx

interne Teilstrecke x zwischen Vermittlungsrechnern



Speichervermittlungsnetz: Innensicht



Vermittlungsrechner

TS-T

Teilstrecke-Teilnehmeranschluss



Puffer für SpV-Datenblock

TSx

interne Teilstrecke x zwischen Speichervermittlungsstellen



SpV-Datenblöcke



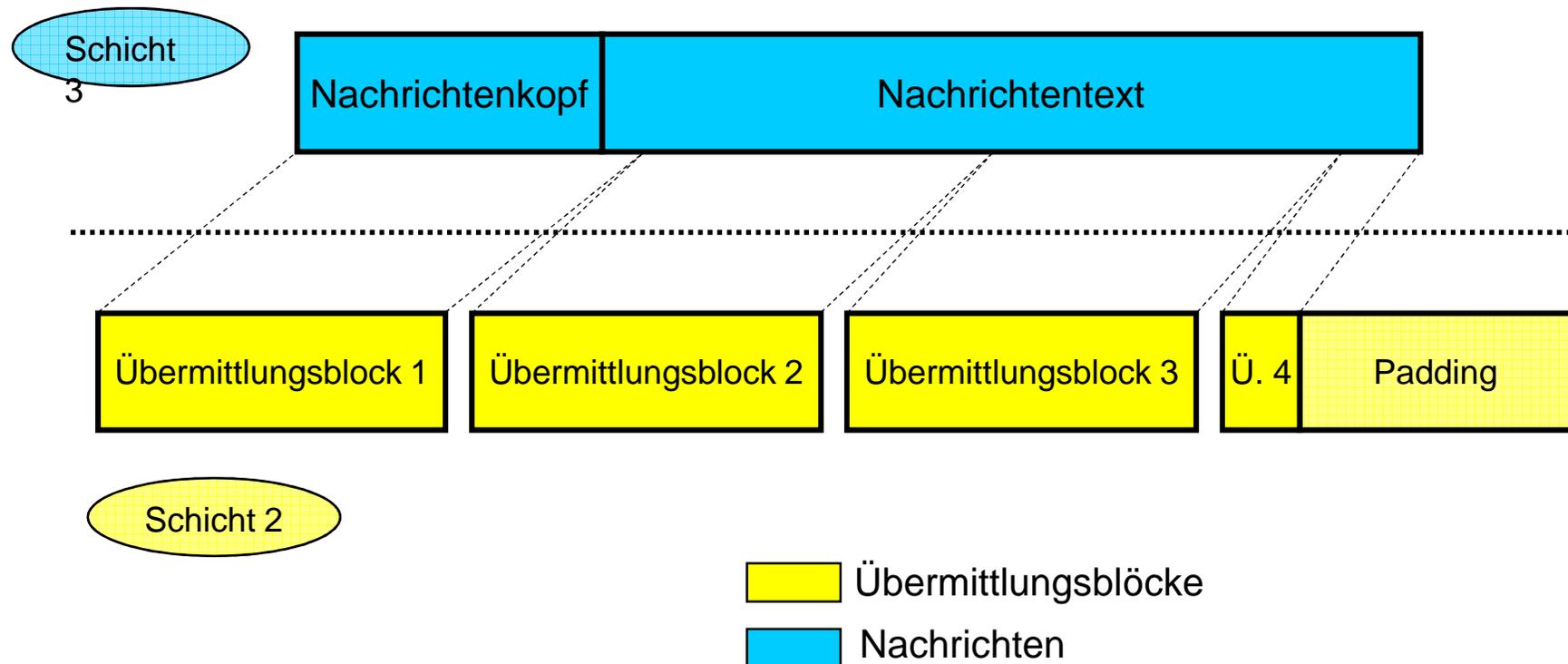
Der Nachrichtenbegriff

- Aufbau einer Nachricht aus
 - Nachrichtenkopf mit Steuerdaten der Schicht 3
 - Ziel- und Herkunftsadresse
 - Nachrichtentyp
 - Längenangabe
 - Priorität
 - Zeitangaben (z.B. Abgangszeit)
 - Nachrichtentext
- Eine Nachricht der Schicht 3 kann in einem oder mehreren Übermittlungsblöcken der Schicht 2 versendet werden.
- Zentrales Kennzeichen
 - Die Nachricht muss beim empfangenen Vermittlungsknoten vollständig eingehen und wieder montiert werden, ehe die Weiterleitung über die nächste Teilstrecke zur folgenden Speichervermittlungsstelle oder zum endgültigen Teilnehmer erfolgen kann.



Segmentieren von Nachrichten

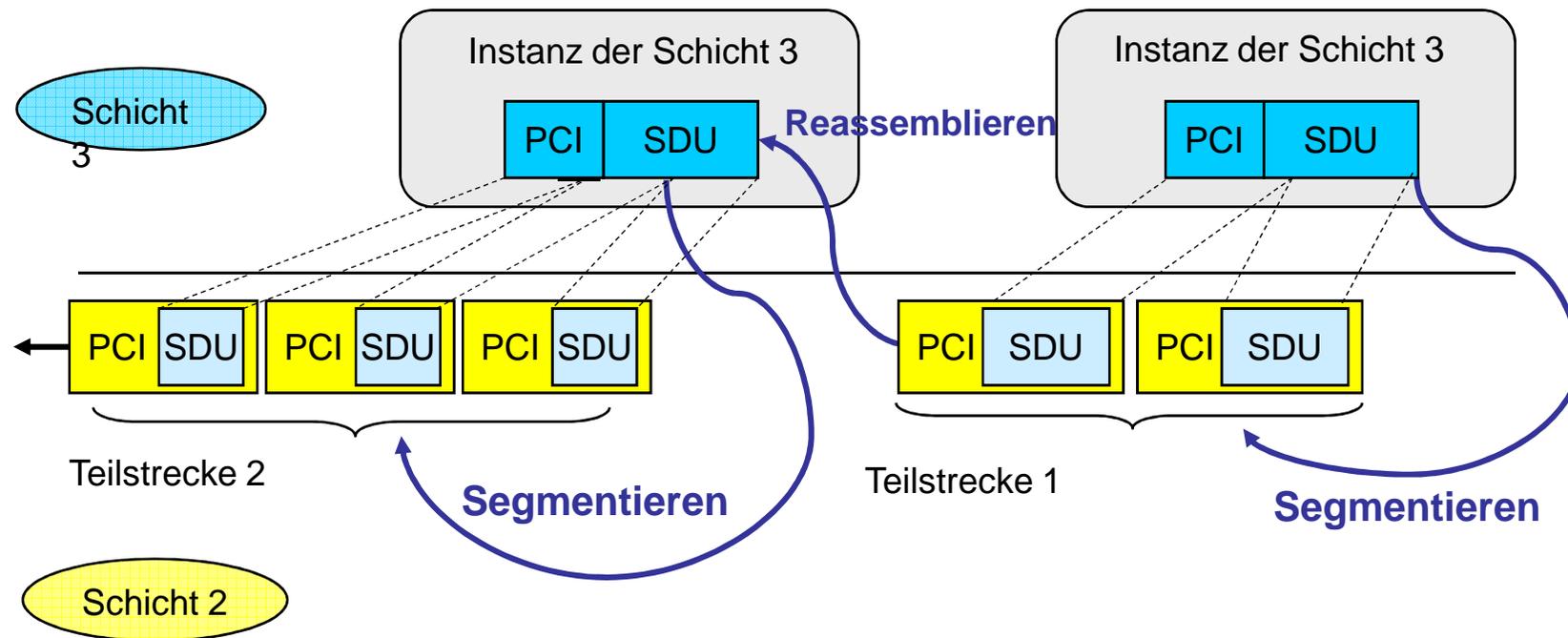
- Segmentieren von Nachrichten (z.B. wegen Blocklängenbegrenzung):
Aufteilen von Nachrichten in mehrere Übermittlungsblöcke
- Resultierende zu kurze Übermittlungsblöcke werden eventuell mit Bits aufgefüllt, um Mindestlänge der Blöcke auf Schicht 2 zu erreichen (Padding)





Segmentieren und Reassemblieren einer Nachricht

- Segmentieren und Reassemblieren von Schicht-3-Nachrichten zur Anpassung an unterschiedliche maximale Blocklängen auf Schicht 2



PCI = Protocol Control Information **der jeweiligen Schicht**

SDU = Service Data Unit **der jeweiligen Schicht**



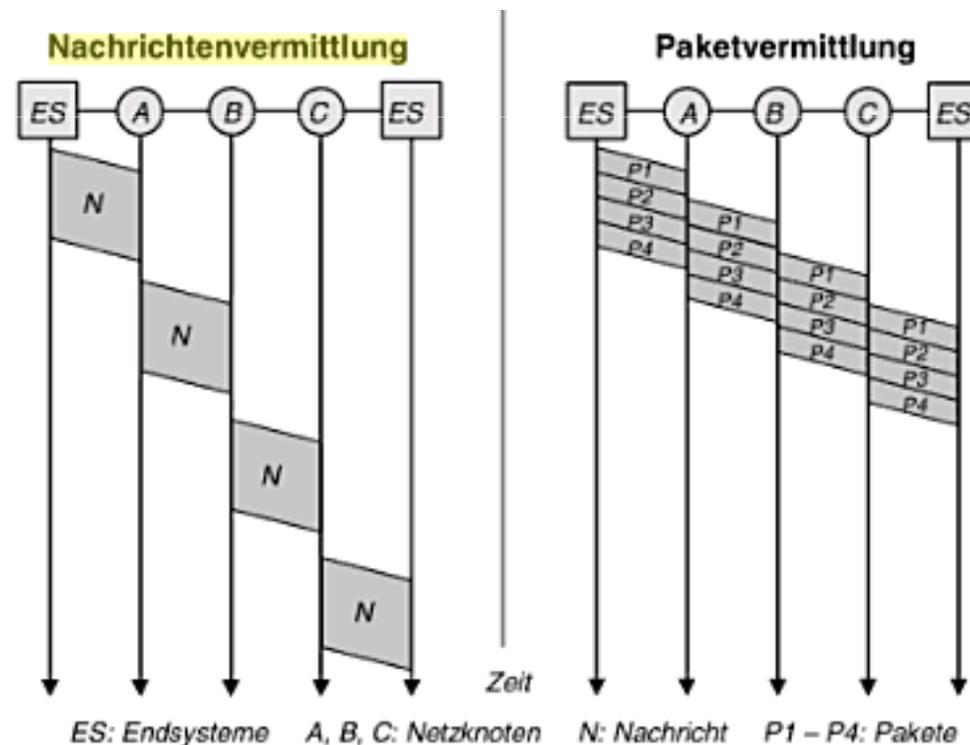
4.2.3 Paketvermittlung

- Paket:
 - vergleichsweise kurzer Datenblock (z.B. einige 100 Oktette oder einige 1000 bit pro Paket)
 - Länge ist fest oder variabel mit vorgegebener Maximallänge
- Vermittlungsvorgang
 - Pakete zeitlich voneinander unabhängig
 - Dadurch zeitlich unabhängige Weitervermittlung je Paket
 - Kein Wiederausammensetzen (Reassemblierung) im Durchgangsvermittlungssystem erforderlich
 - Folge:
Zeitliche Überlappungen möglich, d.h. kürzere Durchlaufzeiten durch das Netz als bei der Nachrichtenvermittlung



Nachrichtenvermittlung vs. Paketvermittlung

- Hauptunterschied zwischen Nachrichten- und Paketvermittlung
 - Paketvermittlung: Inhaltlich zusammengehörende Transfereinheiten (Transport-Datenblöcke der Schicht 4) werden in Pakete nach den Vorschriften des Paketvermittlungsnetzes segmentiert
 - Nachrichtenvermittlung: Wiederherstellung der Transfereinheiten in jedem Vermittlungssystem aus den Segmenten





Paketvermittlung - Vermittlungsklassen

- Vermittlungsklassen für Paketvermittlungsnetze
 - Datagrammvermittlung (Datagram Switching)
 - Stellt eine verbindungslose Vermittlungstechnik dar
 - Virtuelle Verbindung (Virtual Circuit, VC)
 - Feste virtuelle Verbindung (Permanent Virtual Circuit, PVC)
 - Gewählte virtuelle Verbindung (Switched Virtual Circuit, SVC)



Datagrammvermittlung

- Eigenschaften:
 - Pakete bilden geschlossene Vermittlungseinheiten, u.a. mit **vollständigen Adressdaten**:
 - Herkunftsadresse
 - Zieladresse
 - **Kein Verbindungsaufbau** oder sonstige Abstimmungsprozedur über das Netzwerk vor Senden eines Datagrammes notwendig
 - Datagrammvermittlungsstelle ohne Verbindungskontext

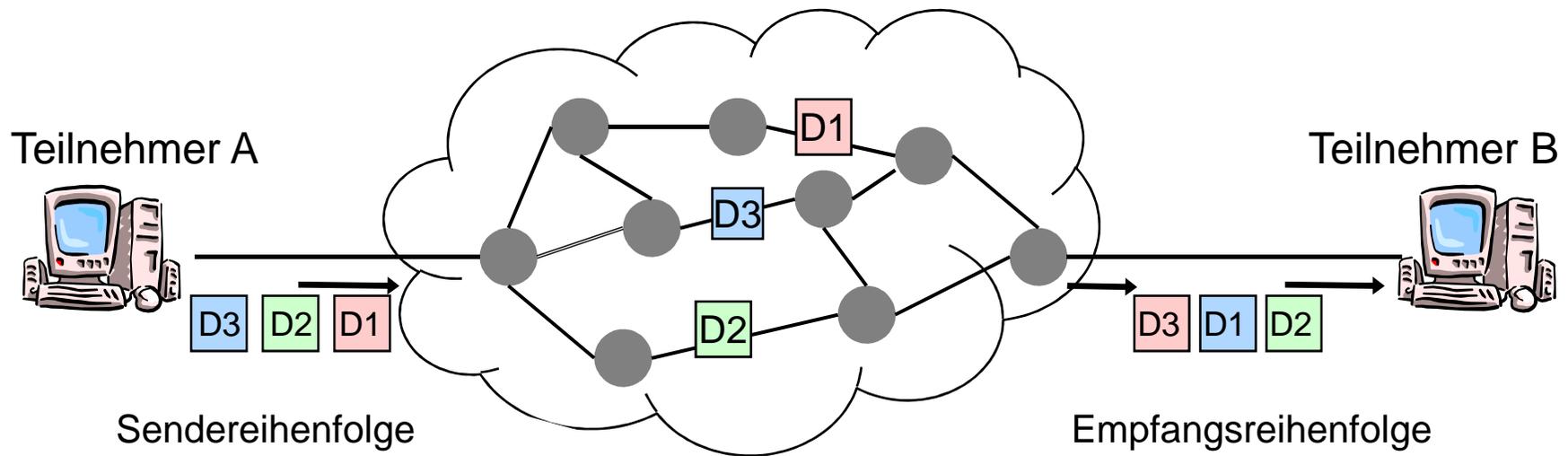
- Vorteile
 - Verbindungsaufbau, -überwachung und -abbau entfallen
 - Bessere Nutzung der Netzkapazität möglich

- Nachteile:
 - Bei vermaschten Netzen mit alternativen Kommunikationswegen können Datagramme zwischen identischen Sendern und Empfängern unterschiedliche Wege zurücklegen
 - **Überholvorgänge möglich**, keine reihenfolgerichtige Auslieferung gewährleistet



Datagrammvermittlung im vermaschten Netz

- Datagrammverkehr in einem Paketvermittlungsnetz
 - Jedes Datagramm mit eigenem Weg

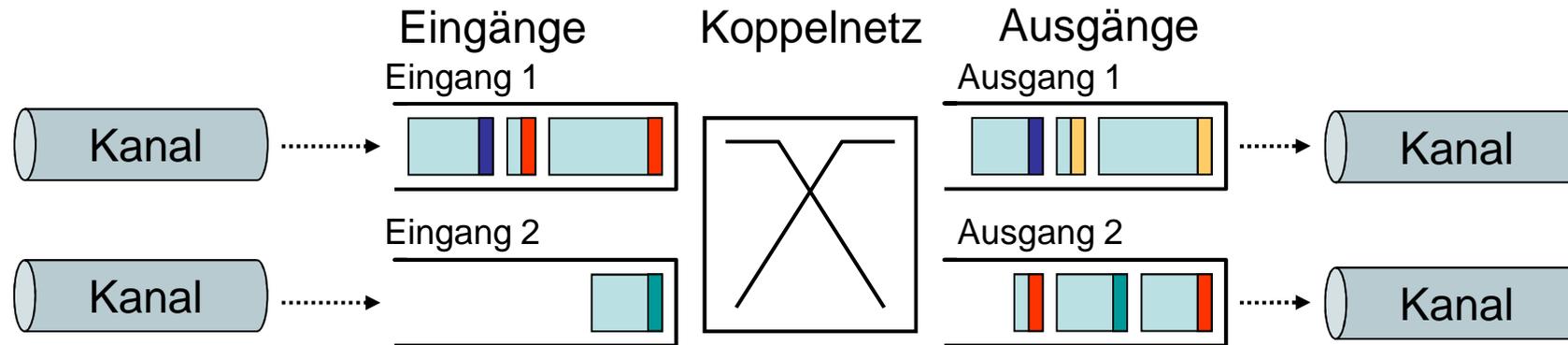


D_x Datagramm *x*

● Datagrammvermittlungsstelle



Datagrammvermittlungsknoten



- Routing-Tabelle:

| Ziel | Ausgang | Next-Hop |
|------|---------|----------|
| A | 1 | X |
| B | 1 | X |
| C | 2 | Y |
| D | 2 | Z |
| ... | ... | ... |

- Für jedes Datagramm wird individuelle Routing-Entscheidung getroffen
- Routing-Entscheidung anhand der Zieladresse im Paketkopf



Virtuelle Verbindungen

- Eigenschaften
 - Bidirektionaler fester Übertragungsweg (vollduplex)
 - Zwischen einem Paar logischer Anschlusspunkte in zwei kommunizierenden Knoten (auf Schicht 3) definiert
 - Kann mehrere Paketvermittlungsstellen umfassen

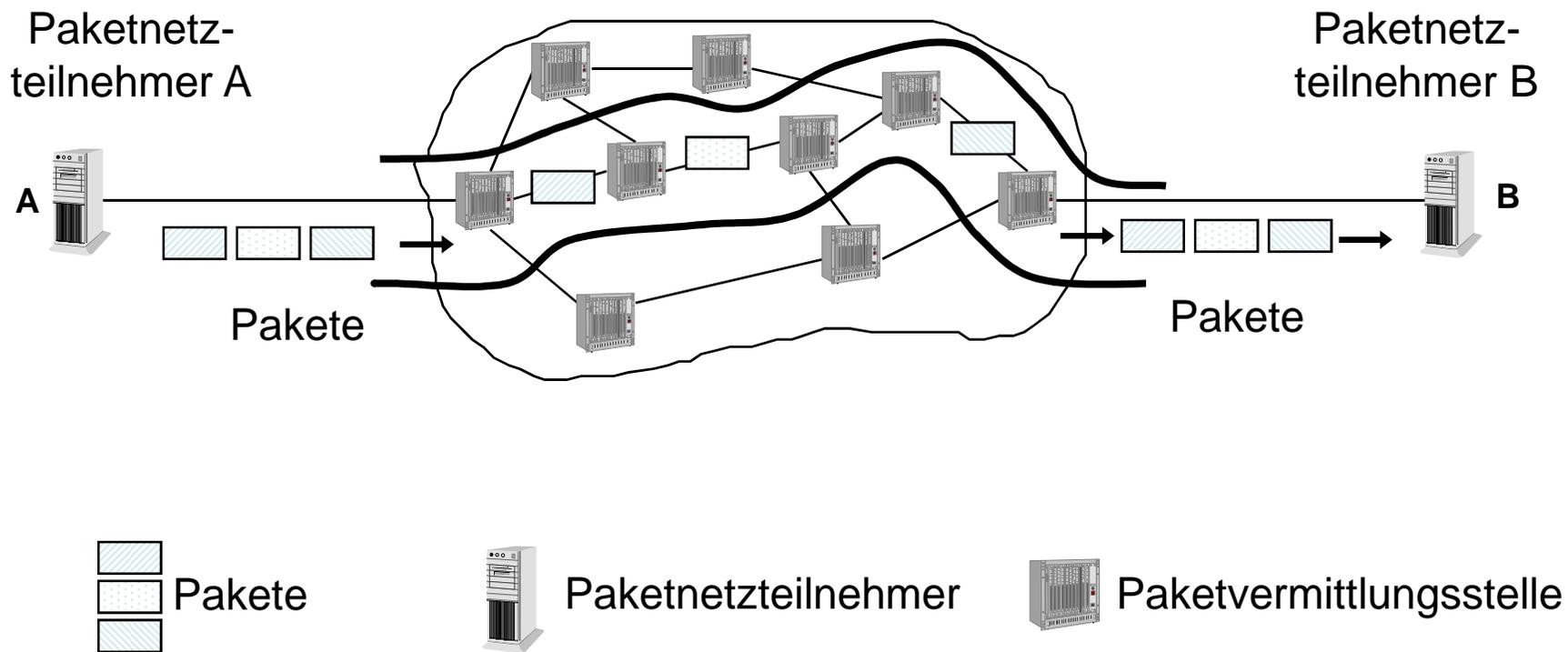
- Feste virtuelle Verbindung
 - Längerfristig eingerichteter Übertragungsweg durch das Netz
 - Ähnlich einer Standleitung in Durchschaltnetzen

- Gewählte virtuelle Verbindung
 - Verbindungsaufbauprozedur erforderlich
 - Aufbau erfolgt vor der Datenaustauschphase
 - Ähnlich wie bei der Leitungsvermittlung



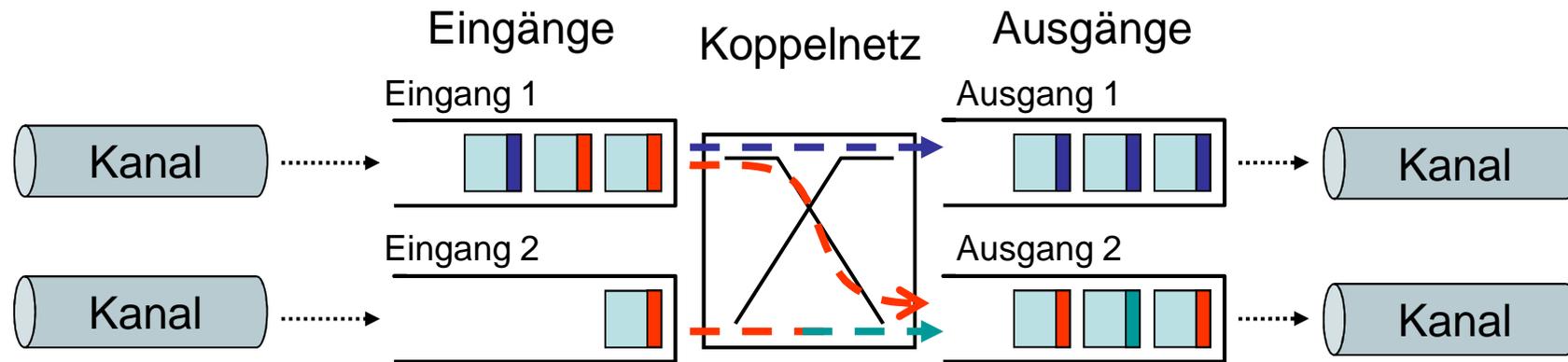
Virtuelle Verbindungen - Verbindungsherstellung

- Schalten einer virtuellen Verbindung über eine feste Route durch die Anwendung des Prinzips der Speichervermittlung
- Alle Pakete verwenden einen identischen Weg im Netz
- Verbindungskontext in beteiligten Paketvermittlungsstellen





Vermittlungsknoten für virtuelle Verbindungen



- Verbindungskontext gespeichert in Weiterleitungstabellen

Eingang 1:

| Eing.-VCI | Ausgang | Ausg.-VCI |
|-----------|---------|-----------|
| A | 1 | A |
| B | 2 | B |
| ... | ... | ... |

Eingang 2:

| Eing.-VCI | Ausgang | Ausg.-VCI |
|-----------|---------|-----------|
| B | 2 | C |
| ... | ... | ... |

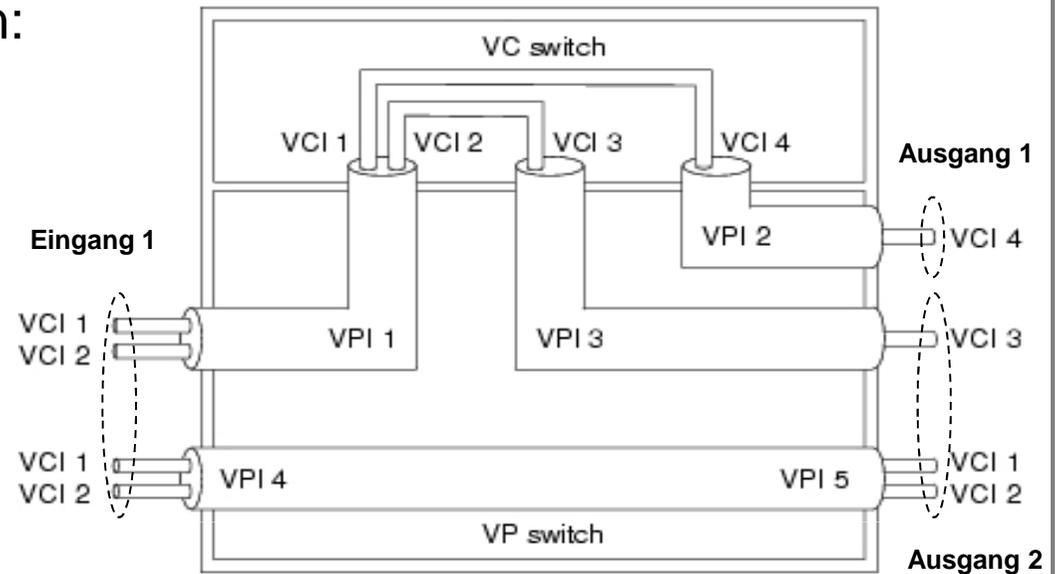
VCI geändert, um Kollision zu vermeiden
⇒ Label Swapping

- Weiterleitungsentscheidung wird anhand eines VCI (Virtual Circuit Identifier) getroffen
- Virtuelle Verbindungen müssen vorher aufgebaut werden



Beispiel: Virtual Circuit Switching bei ATM

- Vermittlung auf zwei Ebenen:
 - Virtual Channel
= einzelne Verbindung
 - Virtual Path
= Aggregat mehrerer VCs

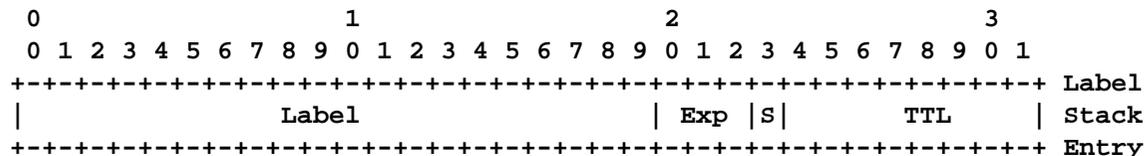
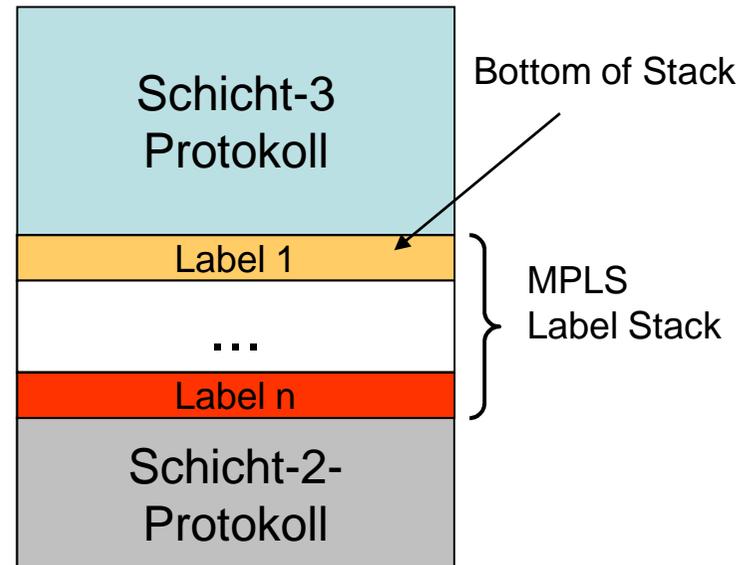


| Eing.-VPI / VCI | Ausgang | Ausg.-VPI / VCI |
|-----------------|---------|-----------------|
| 1 / 1 | 1 | 2 / 4 |
| 1 / 1 | 2 | 3 / 3 |
| 4 / * | 2 | 5 / * |
| ... | ... | ... |



Beispiel: Multi-Protocol Label Switching (MPLS)

- Eigenschaften:
 - virtuelle Verbindung für beliebige Technologien und Protokolle
 - beliebig viele Labels im Stack
- Verarbeitung der Labels in LSRs (Label Switched Routers):
 - Hinzufügen und Entfernen von Labels
 - label-abhängige Weiterleitung
- Label Distribution Protocol (LDP):
 - Signalisierung und Label-Austausch zwischen LSRs



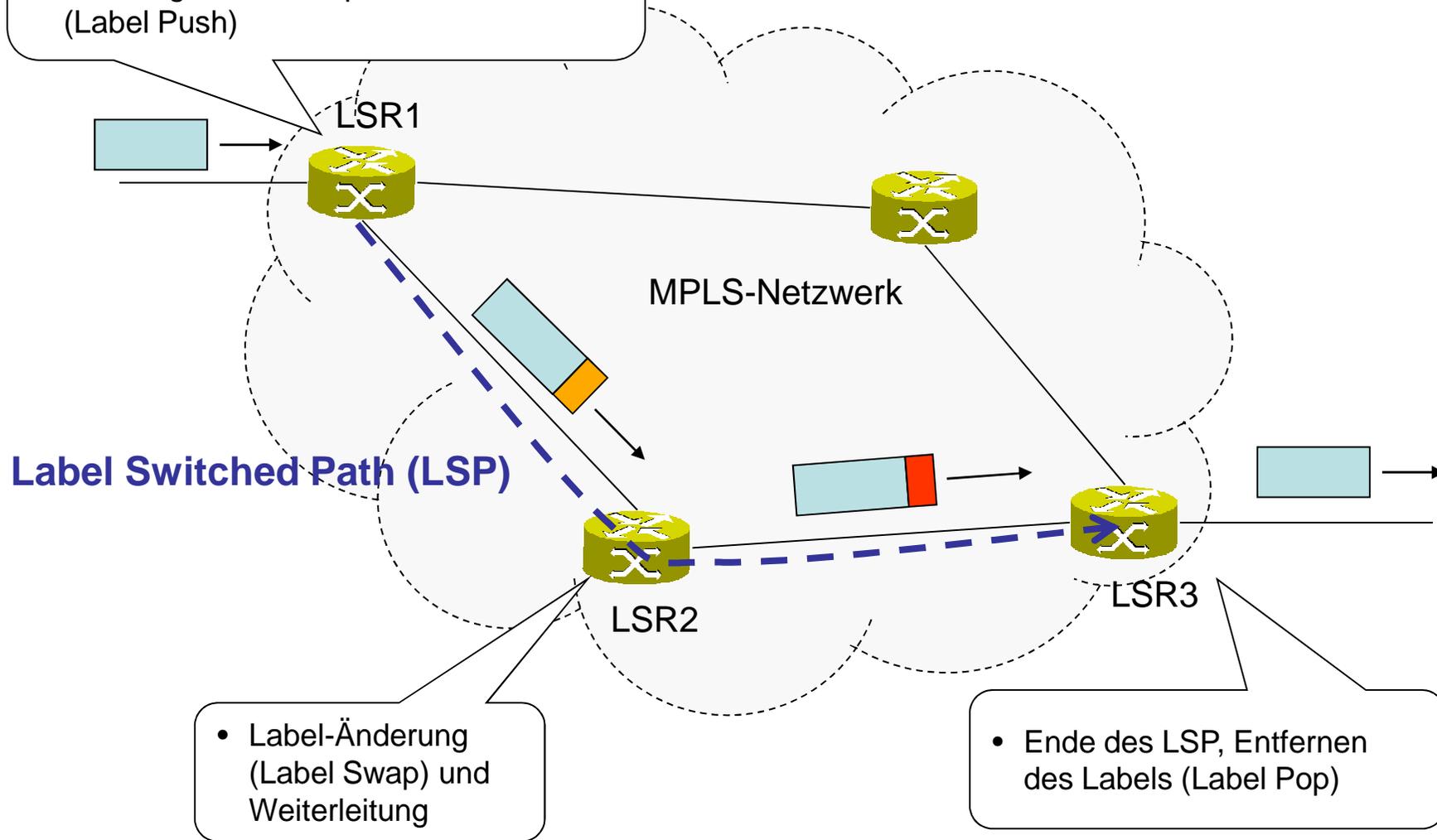
Label: Label Value, 20 bits
 Exp: Experimental Use, 3 bits
 S: Bottom of Stack, 1 bit
 TTL: Time to Live, 8 bits



MPLS: Label Switched Path

- Paketeinteilung in FECs (Forward Equivalent Classes)
- Hinzufügen des entsprechenden Labels (Label Push)

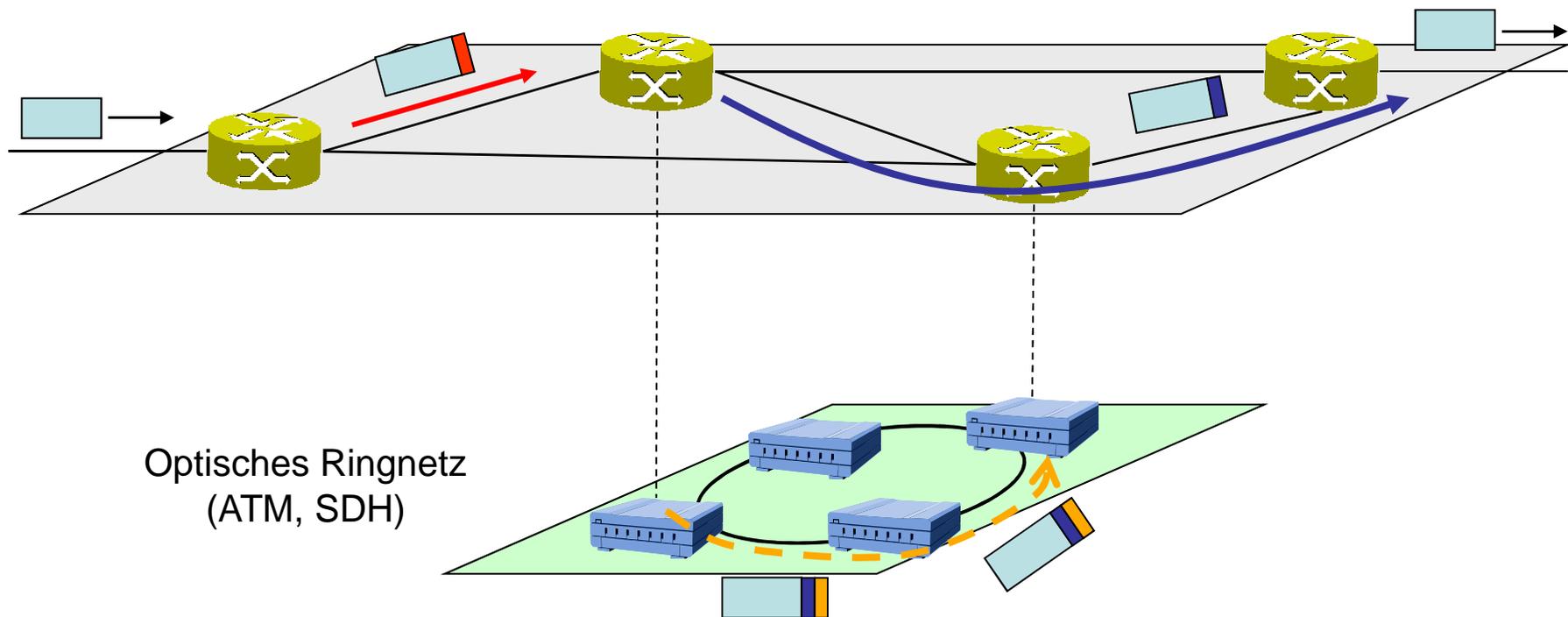
FEC: Menge von Paketen, die im MPLS-Netz gleich behandelt werden sollen (gleiche Route, gleiche QoS)





MPLS-Hierarchie

- Label-Switching auf mehreren Schichten





MPLS-Anwendungen

- Eine Motivation für virtuelle Verbindungen war:
 - größerer Durchsatz dank einfachen Forwarding-Entscheidungen in den Vermittlungsknoten/Routern
 - ➔ heute kein Argument mehr, weil Datagramm-Vermittlung sehr schnell und effizient in Hardware implementiert werden kann

- Anwendungsgebiete von MPLS:
 - Traffic Engineering: gezieltes Routing einzelner Verkehrsströme
 - Quality of Service: unterschiedliche Paketbehandlung und –weiterleitung je nach Dienstgüteanforderung
 - VPN (Virtual Privat Networks)

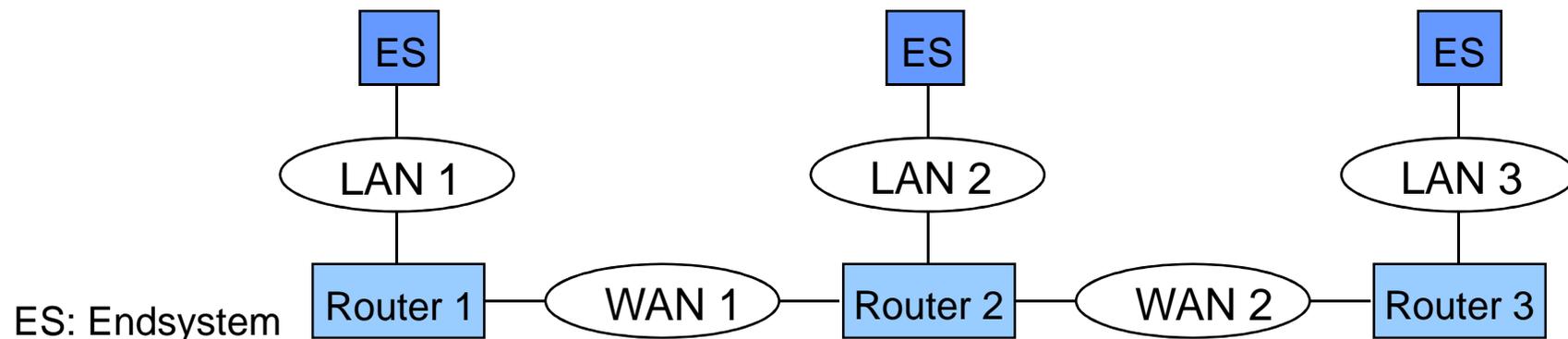
- GMPLS (Generalized MPLS):
 - Ausweitung auf Transportnetze mit physikalischer Leitungsvermittlung (v.a. optische Netze)
 - Label entspricht einem Zeitschlitz bei TDM oder einer Wellenlänge bei WDM (Wavelength Division Multiplex)



4.2.4. Router - Kopplung von Netzwerken auf Schicht 3

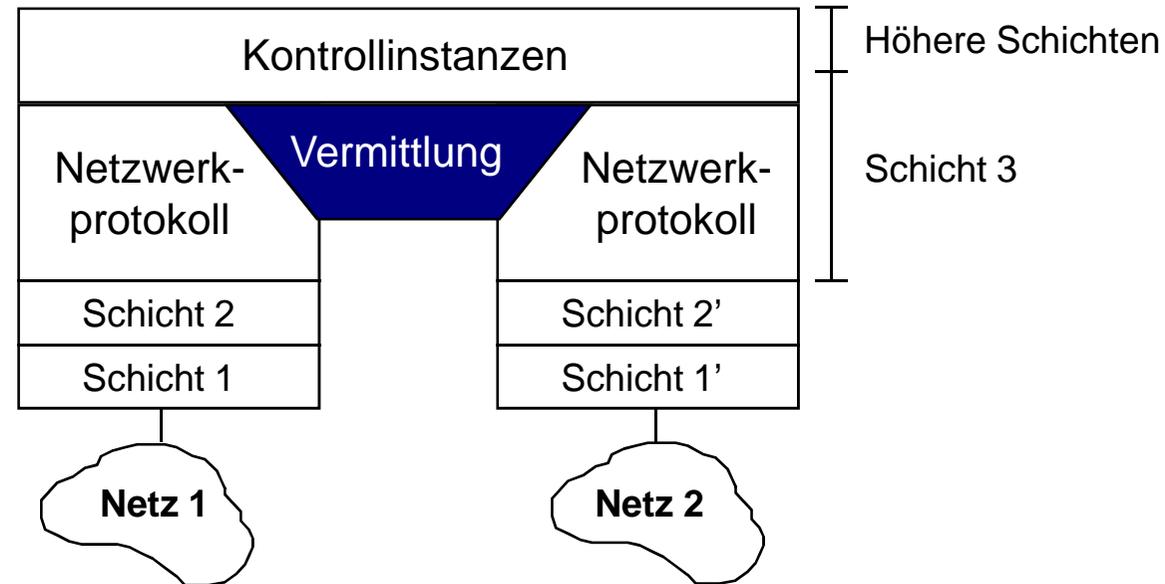
□ Aufgaben:

- Kommunikation entfernter Endsysteme über ein oder mehrere WANs
- Wegewahl anhand weltweit eindeutiger, bevorzugt hierarchischer Netzwerkadressen (z.B. IP-Adressen)
- Segmentieren/Reassemblieren von Schicht-3-Datenpaketen zur Anpassung an unterschiedliche maximale Paketgrößen auf Schicht 2
- Sicherheitsmechanismen zur Regelung von Netzzugriffen abhängig von der Netzwerkadresse (Stichwort „Firewall“)
- Automatische Begrenzung von Schicht-2-Broadcasts





Architektur eines Routers

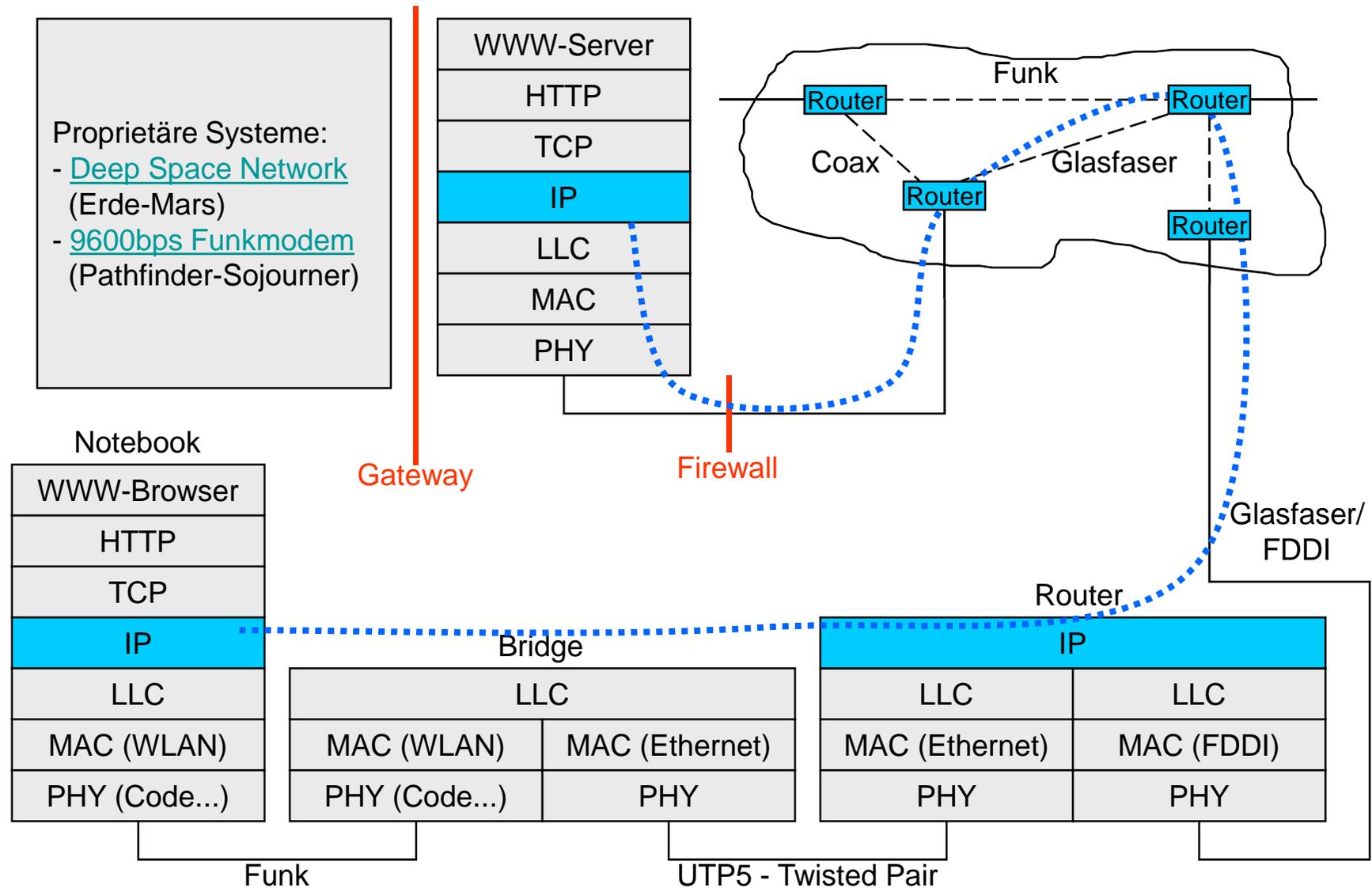


□ Wesentliche Merkmale:

- Für jedes Netzwerk eine eigene Schicht-1- und Schicht-2-Instanz
- Netzwerkprotokoll ist in der Regel für alle Netzwerke gleich (z.B. IP-Router)
- Wegwahl anhand der global eindeutigen Netzwerkadressen
- Vermittlungskomponente verbindet die Netzwerkprotokollinstanzen; sie realisiert die Weiterleitungsfunktion
- Kontrollinstanzen implementieren beispielsweise Routing-Protokolle, Protokolle zur Fehleranzeige und Managementprotokolle



Routing im Beispiel





Wegewahl vom Notebook zur NASA

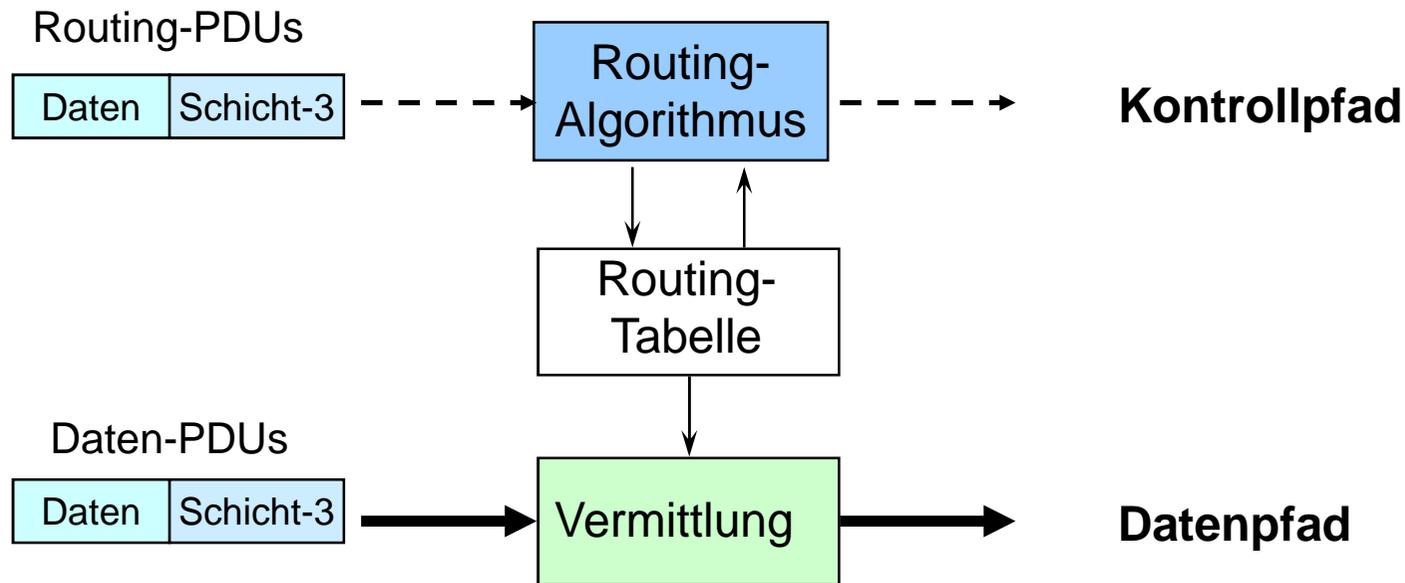
Aktueller Weg kann z.B. im Internet mit Hilfe von traceroute angezeigt werden

```
traceroute to www.NASA.GOV (128.183.243.3)
```

| HOP | NAME (IP-address) | TIME | probe 1 | probe 2 | probe 3 |
|-----|---|------|---------|---------|---------|
| 0 | mobile1.telematik.informatik.uni-karlsruhe.de (129.13.35.123) | | | | |
| 1 | i70route1 (129.13.35.244) | | 9 ms | 9 ms | 10 ms |
| 2 | iracs1.ira.uka.de (129.13.1.1) | | 2 ms | 3 ms | 2 ms |
| 3 | Karlsruhel.BelWue.DE (129.143.59.1) | | 2 ms | 3 ms | 2 ms |
| 4 | Uni-Karlsruhel.WiN-IP.DFN.DE (188.1.5.29) | | 3 ms | 3 ms | 3 ms |
| 5 | ZR-Karlsruhel.WiN-IP.DFN.DE (188.1.5.25) | | 5 ms | 3 ms | 2 ms |
| 6 | ZR-Frankfurt1.WiN-IP.DFN.DE (188.1.144.37) | | 8 ms | 7 ms | 7 ms |
| 7 | IR-Frankfurt1.WiN-IP.DFN.DE (188.1.144.97) | | 8 ms | 11 ms | 9 ms |
| 8 | IR-Perryman1.WiN-IP.DFN.DE (188.1.144.86) | | 124 ms | 126 ms | 102 ms |
| 9 | border3.Washington.mci.net (166.48.41.249) | | 121 ms | 123 ms | 124 ms |
| 10 | core4.Washington.mci.net (204.70.4.105) | | 123 ms | 135 ms | 121 ms |
| 11 | mae-east4.Washington.mci.net (204.70.1.18) | | 123 ms | 122 ms | 121 ms |
| 12 | mae-east.nsn.nasa.gov (192.41.177.125) | | 125 ms | 126 ms | 126 ms |
| 13 | rtr-wan1-ef.gsfc.nasa.gov (192.43.240.33) | | 127 ms | 123 ms | 121 ms |
| 14 | rtr-600.gsfc.nasa.gov (128.183.251.26) | | 147 ms | 140 ms | 130 ms |
| 15 | bolero.gsfc.nasa.gov (128.183.243.3) | | 127 ms | 133 ms | 129 ms |



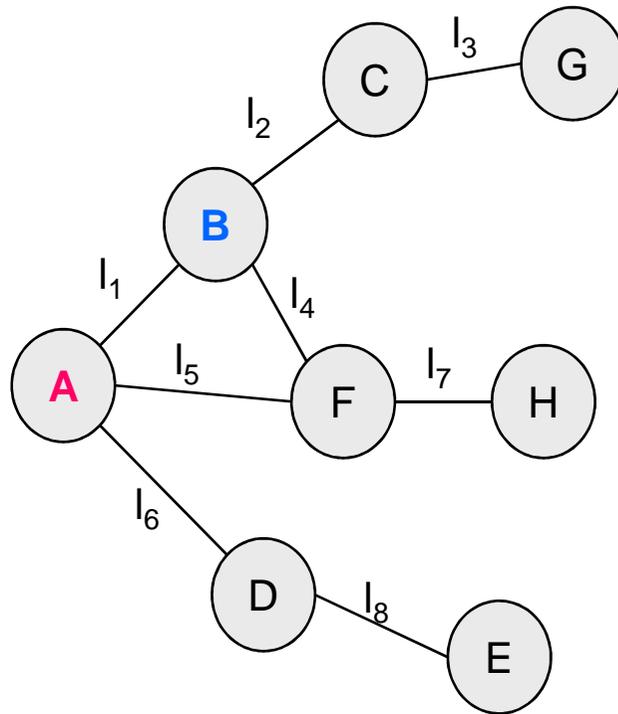
Router: Kontroll- und Datenpfad



- Datenpfad auf Netzwerkschicht
- Kontrollpfad darüber (Routing-PDUs sind in N-PDUs oder T-PDUs gekapselt)
- Gewinnung von Routinginformationen durch das **Routing-Protokoll**
- **Routing-Algorithmus** verwaltet die Routing-Tabelle bzw. Forwarding-Tabelle (Einfügen/Löschen/Ändern von Einträgen) auf der Basis der gewonnenen Routinginformation
- **Routing-Tabelle** bzw. Forwarding-Tabelle enthält Routinginformationen
- Wegewahl bei der Vermittlung wird anhand der Routing-Tabelle bzw. Forwarding-Tabelle durchgeführt



Prinzip einer Routingtabelle



A-H: Vermittlungsrechner (Router)
l₁-l₈: Abgehende/ankommende Teilstrecken (Links)

Routingtabelle
in **Router B**

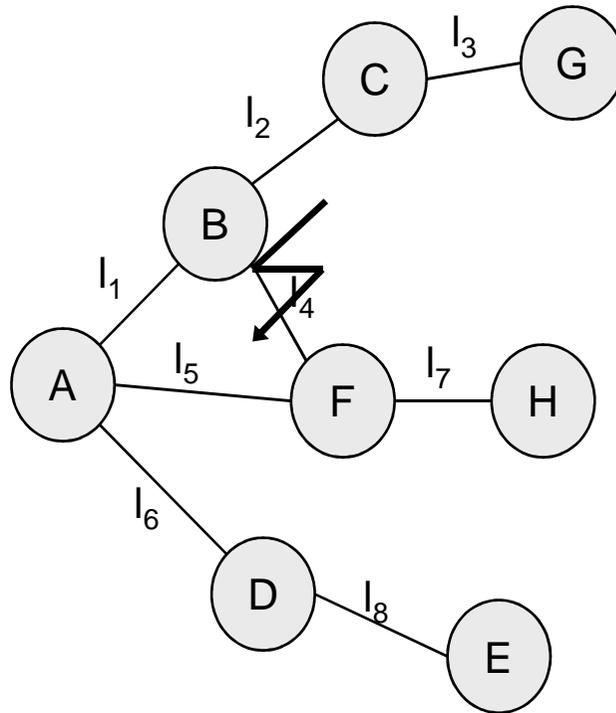
| nach | über |
|------|---------------------|
| A | - (l ₁) |
| C | - (l ₂) |
| D | A (l ₁) |
| E | A (l ₁) |
| F | - (l ₄) |
| G | C (l ₂) |
| H | F (l ₄) |

Routingtabelle
in **Router A**

| nach | über |
|------|------|
| B | |
| C | |
| D | |
| E | |
| F | |
| G | |
| H | |



Prinzip einer Routingtabelle: Ausfall eines Links



A-H: Vermittlungsrechner (Router)
l₁-l₈: Abgehende/ankommende
Links

Routingtabelle
in Router B

| nach | über |
|------|---------------------|
| A | - (l ₁) |
| C | - (l ₂) |
| D | A (l ₁) |
| E | A (l ₁) |
| F | A (l ₁) |
| G | C (l ₂) |
| H | A (l ₁) |

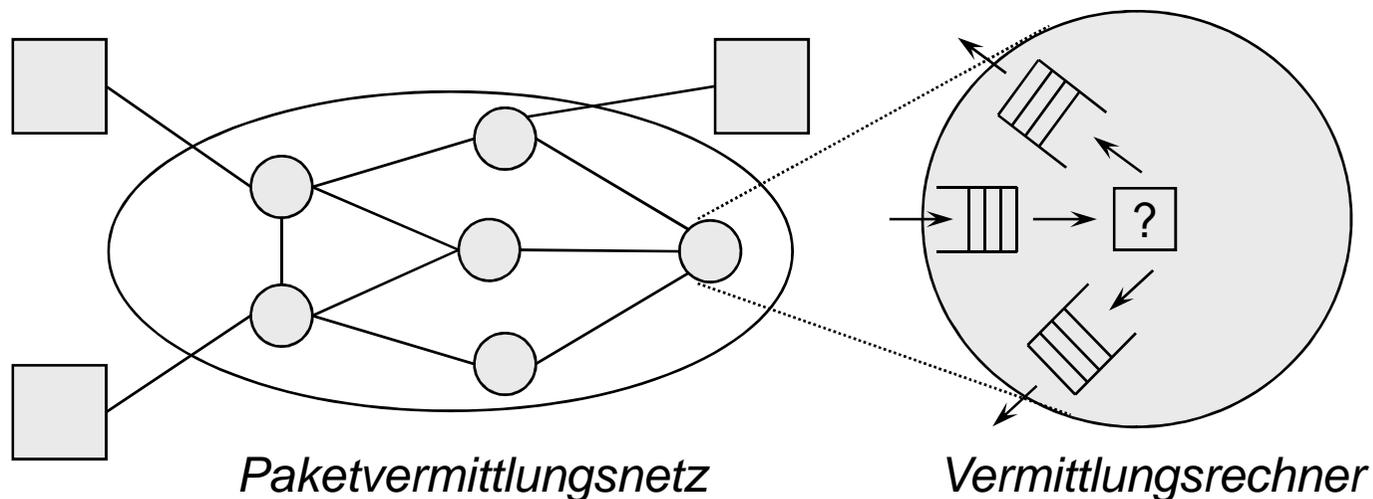
Routingtabelle
in Router A

| nach | über |
|------|------|
| B | |
| C | |
| D | |
| E | |
| F | |
| G | |
| H | |



4.2.5. Routingverfahren

- Aufgaben
 - Fällen der Entscheidung, auf welcher Übertragungsleitung ein eingehendes Paket (Nachricht) weitergeleitet werden soll
- Ziele
 - Niedrige mittlere Paketverzögerung
 - Hoher Netzdurchsatz
- Ansatzpunkt
 - Übertragen eines Pakets von einem Quellrechner zu einem Zielrechner über einen Weg mit geringsten "Kosten".





Routingverfahren - Routingnetzwerk und Wegewahl

- Definition des Routingnetzwerkes
 - Welche Knoten haben Vermittlungsfunktion?
 - Wie sind die Vermittlungsknoten verbunden: durch einfache oder mehrfache Verbindungen ?

- Weggenerierung
 - Welche Informationen sind zur Weggenerierung notwendig ?
 - Welche Parameter beeinflussen die Kostenfunktion ?
 - Feste Kosten für jede Verbindung (i.allg. umgekehrt proportional der Übertragungskapazität)
 - Anzahl der auf Übertragung wartenden Pakete
 - Fehlerrate
 - Paketverzögerungszeit auf einer Verbindung
 - Art des Verkehrs (Dialog, Batch)

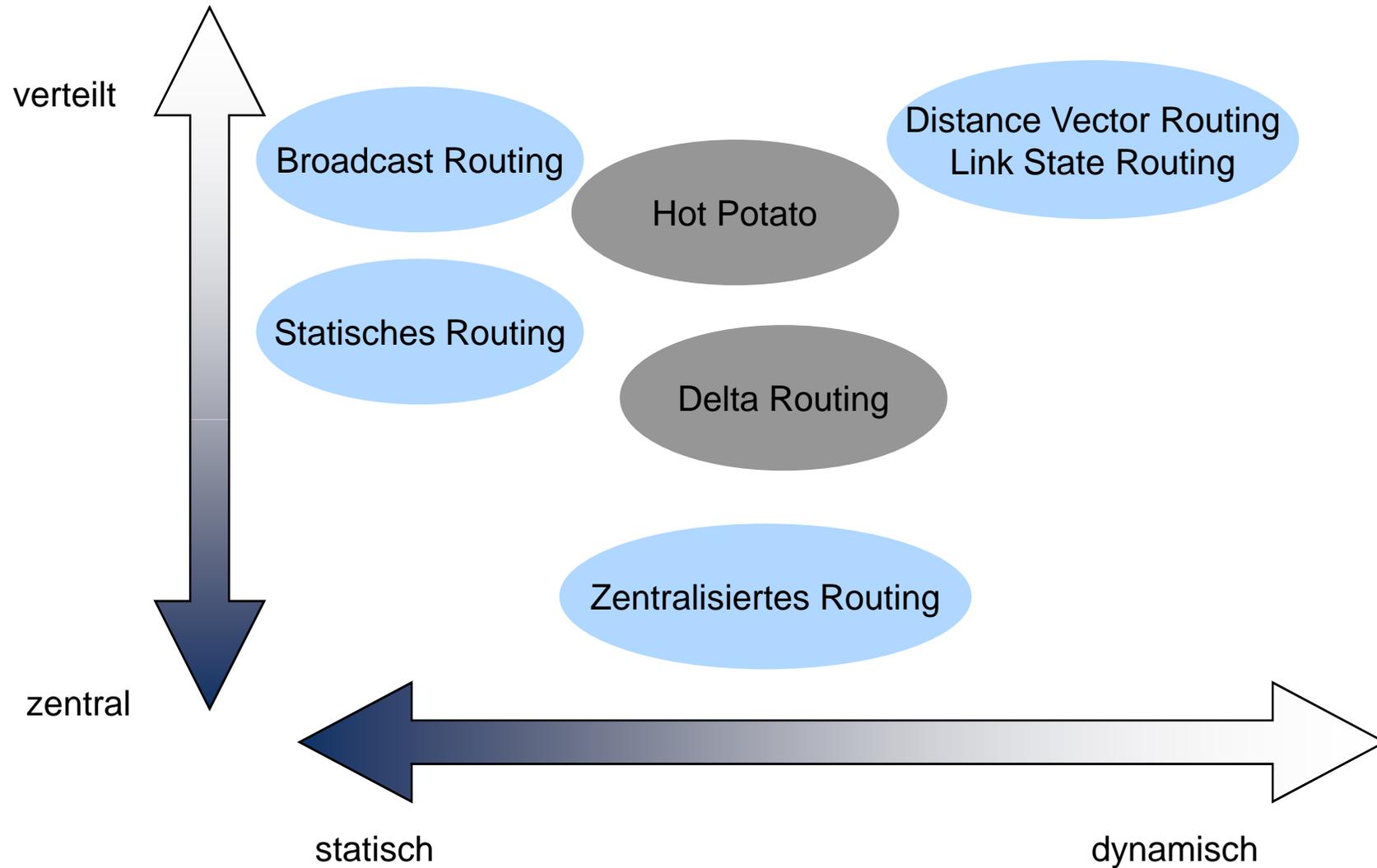


Routingverfahren - Zentralisation und Dynamik

- Zentralisation
 - Wo ist der Routingalgorithmus lokalisiert?
 - Zentral (in einem Netzkontrollzentrum)
 - Dezentral (verteilt auf die Vermittlungsknoten)
- Wie dynamisch ist das Routingverfahren?
 - Nicht adaptiv: Die Routingtabellen in den Vermittlungsknoten bleiben über längere Zeit konstant, verglichen mit Verkehrsänderungen.
 - Adaptiv: Routing-Entscheidungen hängen vom Zustand des Netzes ab (Topologie, Lastverhältnisse).
- Zielkonflikt
 - Knoten haben veraltete oder unvollständige Informationen über den Zustand des Netzes.
 - Belastung durch Austausch von Routinginformationen



Routingverfahren im Überblick





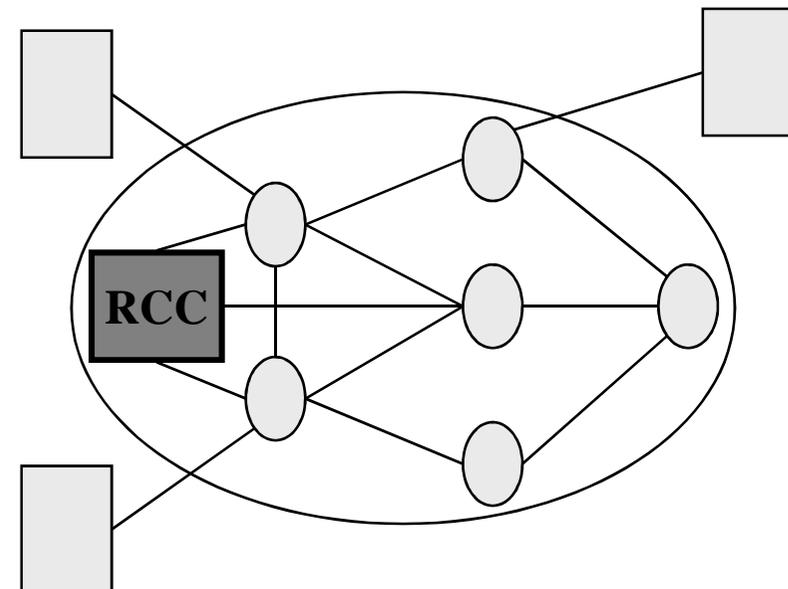
Statisches Routing mit Lastverteilung

- Statisches Routing (Static Routing, Directory Routing) mit Lastverteilung
 - Nicht adaptiv, einfach, viel benutzt
 - Jeder Knoten unterhält eine Tabelle mit einer Zeile für jedes mögliche Zielnetz.
 - Eine Zeile enthält n Einträge, welche die beste, zweitbeste, etc. Übertragungsleitung für dieses Ziel, zusammen mit einer relativen Gewichtung, angeben.
 - Vor der Weiterleitung eines Pakets wird eine Zufallszahl gezogen und eine der Alternativen anhand der Gewichtung ausgewählt.



Zentralisiertes Routing

- Adaptives Verfahren (Delta Routing)
 - Im Netz gibt es ein Routing Control Center (RCC).
 - Jeder Knoten sendet periodisch Zustandsinformationen an das RCC, z.B.:
 - Liste aller aktiven Nachbarn
 - Aktuelle Warteschlangenlängen
 - Umfang an Verkehr, der seit dem letzten Bericht abgewickelt wurde
 - Das RCC sammelt diese Zustandsinformationen und berechnet aufgrund dieser Kenntnis über das gesamte Netz die optimalen Wege zwischen allen Knoten (z.B. kürzeste Wege).
 - Jeder Knoten trifft seine Routing-Entscheidungen anhand der ihm zugewiesenen Routing-Tabelle.





Zentralisiertes Routing - Vor- und Nachteile

- Vorteile
 - Das RCC hat theoretisch die vollständige Übersicht und kann perfekte Entscheidungen treffen.
 - Knoten müssen keine aufwendigen Routing-Berechnungen durchführen.

- Nachteile
 - Für große Netze dauert die Berechnung u.U. sehr lange.
 - Ausfall des RCC lähmt das ganze Netz (Back-up Rechner).
 - Globale Inkonsistenzen möglich, da Knoten nahe dem RCC neue Routing-Tabellen ggf. früher erhalten als die weiter entfernten.
 - Belastung des RCC durch die zentrale Funktion



Isoliertes Routing: Überblick

- Isoliertes Routing (Isolated Routing)
 - Jeder Knoten entscheidet nur aufgrund der Information, die er selbst sammelt.
 - Kein Austausch von Routing-Informationen zwischen den Knoten.
 - Anpassung an Verkehrs- und Topologieänderungen kann damit nur mit Hilfe beschränkter Informationen erfolgen.
 - Unterschiedliche Verfahren
 - Broadcast Routing
 - Backward Learning
 - Hot Potato (Deflection Routing)
 - ...



Broadcast-Routing

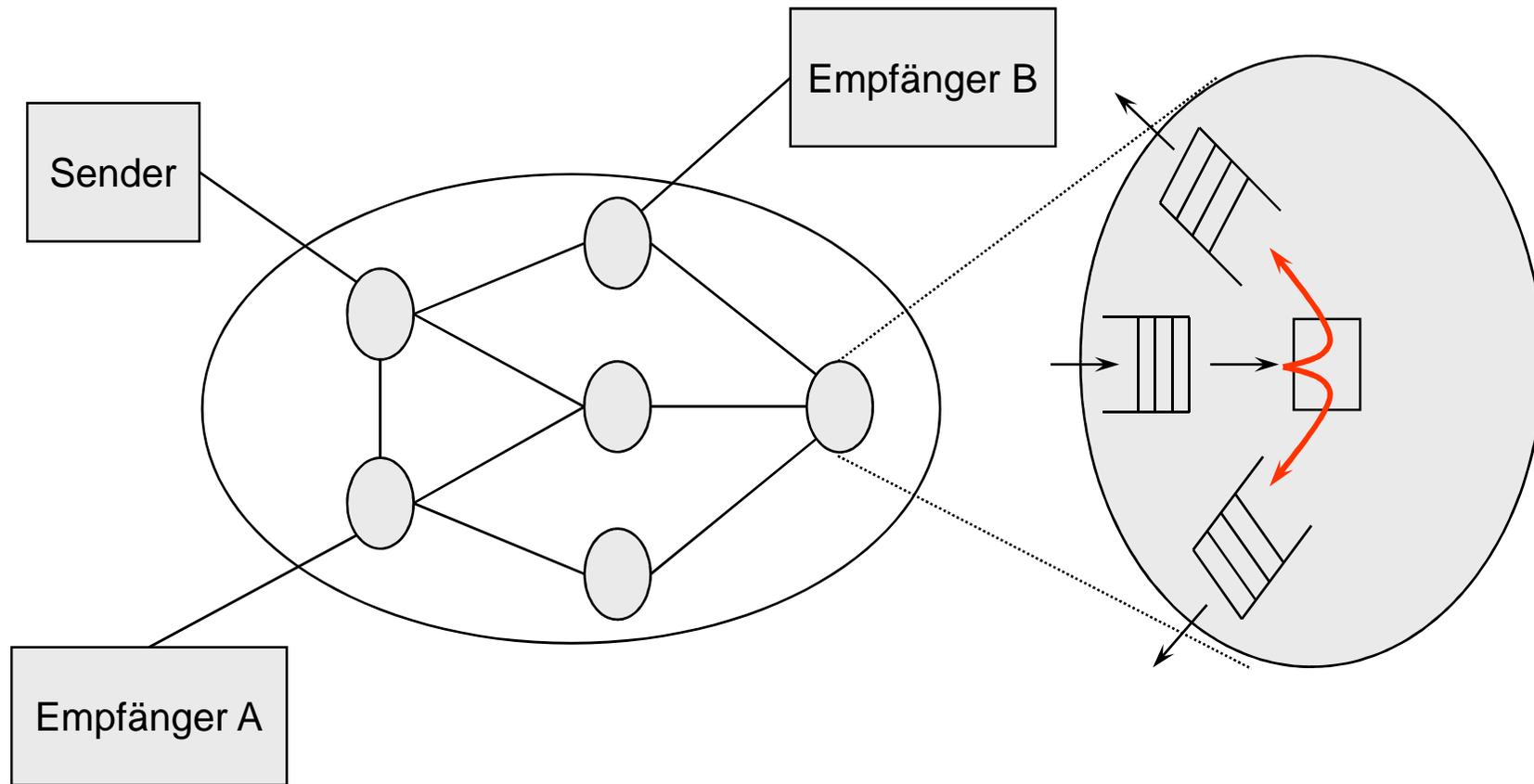
- Broadcast-Routing
 - Senden eines Pakets an alle Knoten, z.B. für Umfragen nach bestimmten Betriebsmittel

- Varianten
 - Erstellen eines gesonderten Paketes für jeden Knoten
 - Fluten

- Arten
 - Multidestination Routing
 - Reverse-Path-Forwarding



Broadcast-Routing: Fluten





Broadcast-Routing: Fluten

- Fluten (Flooding)
 - Einfachstes Verfahren, nicht adaptiv
 - Jedes eingehende Paket wird auf jeder Übertragungsleitung weiter übertragen, außer auf derjenigen, auf der es eintraf.
- Maßnahmen zur Eindämmung der Flut
 - Erkennung von Duplikaten durch die Nummerierung der Pakete
 - Kontrolle der Lebensdauer eines Pakets durch Zählen der zurückgelegten Teilstrecken (hops). Ein hop-Zähler im Paket wird mit der minimalen Zahl von Teilstrecken (idealer Fall) bzw. einer (geschätzten) Zahl für die maximale Zahl von Teilstrecken zwischen Quelle und Ziel initialisiert. In jedem Knoten wird der Zähler um 1 dekrementiert. Falls der Zähler den Wert 0 erreicht, kann das Paket verworfen werden.
 - Varianten:
 - Selektives Fluten
Weiterleitung nicht auf allen, sondern nur auf einigen Leitungen
 - Random Walk
Zufällige Auswahl einer Leitung

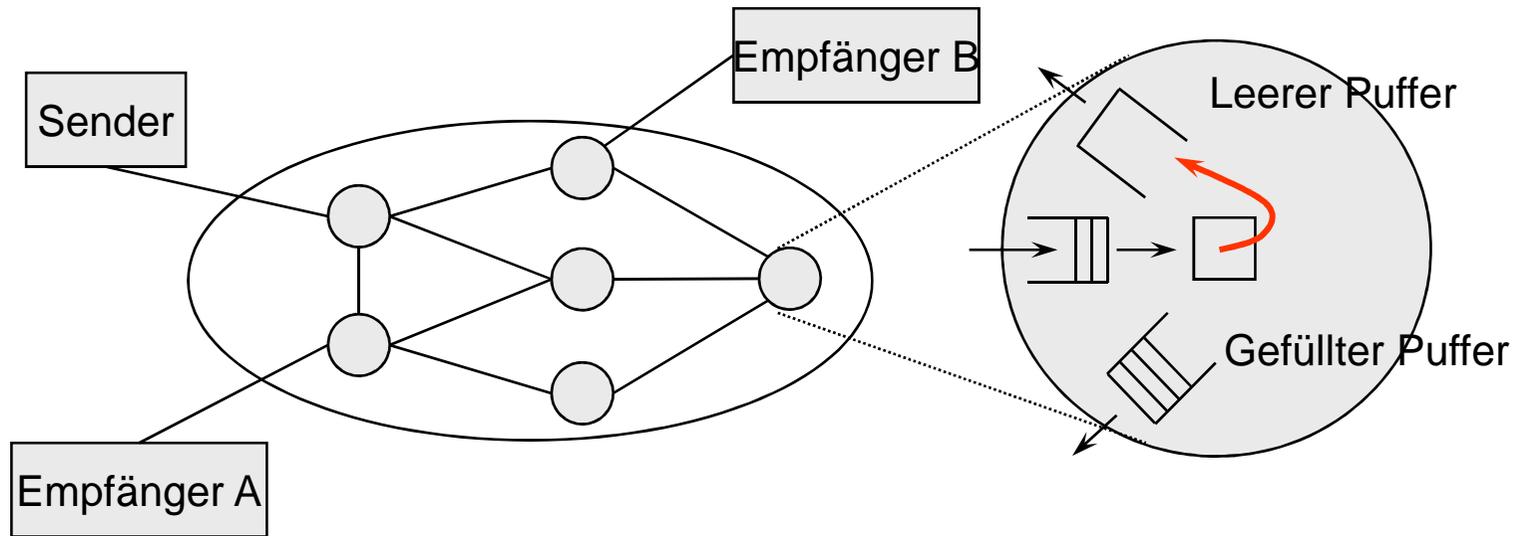


Hot Potato (Deflection Routing)

- Jeder Knoten versucht, eingehende Pakete so schnell wie möglich weiterzuleiten.
- Die Übertragungsleitung mit der kürzesten Warteschlange wird für die Weiterleitung ausgewählt.
- Variante: Kombination mit statischem Routing
 - Auswahl der besten Übertragungsleitung nach statischem Verfahren, solange deren Warteschlangenlänge unter einer bestimmten Schwelle bleibt
 - Auswahl der Übertragungsleitung mit kürzester Warteschlange, falls deren statisches Gewicht nicht zu niedrig ist
- Alternative Bedeutung von "Hot Potato Routing": Jeder Knoten einer administrativen Domäne versucht, ein Paket, das an eine andere administrative Domäne weitergeleitet werden soll, auf dem schnellsten Weg an diese andere Domäne weiterzugeben. (Führt zu asymmetrischen Pfaden mit unterschiedlichem Hin- und Rückweg.)



Hot Potato (Deflection Routing)





Backward Learning

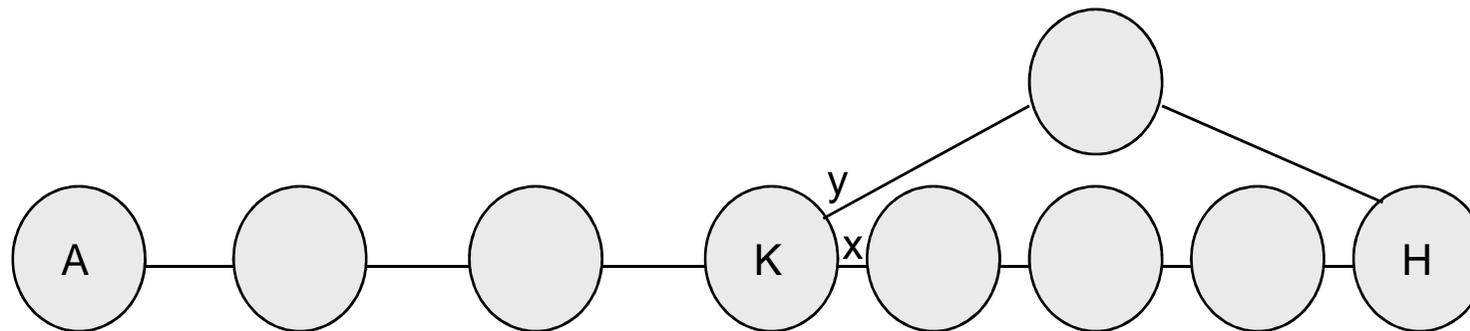
- Daten im Paket
 - Identifikation des Quellknotens
 - Ein Zähler, der mit jeder zurückgelegten Teilstrecke (hop) um 1 erhöht wird.

- Beispiel
 - Falls z.B. bei einem Knoten K auf der Übertragungsleitung k ein Paket mit Zähler = 4 vom Ursprungsknoten H eintrifft, so weiß der Knoten K, dass er den Knoten H über die Leitung k in 4 hops erreichen kann.
 - Falls Knoten Ks bisheriger bekannter optimaler Weg zu Knoten H (über eine andere Leitung) mehr als 4 hops beträgt, so aktualisiert Knoten K seine Routingtabelle mit dem neuen und jetzt besseren Weg.



Backward Learning

| Ziel- adresse | Herkunfts- adresse | Hop- zähler | |
|------------------|-----------------------|----------------|--|
| A | H | | |



| Routingtabelle | Ziel | Ausgang | Hops |
|----------------|------|---------|------|
| K | . | . | . |
| | H | x | 4 |



Backward Learning - Nachteile und Probleme

- Nachteile
 - Nur Änderungen zum Besseren werden zur Kenntnis genommen
 - Ausfälle oder Überlastung von Übertragungsleitungen werden nicht weitergemeldet.

- Folge
 - Knoten müssen periodisch alle Informationen vergessen und wieder initial aufsetzen.

- Problem
 - Während der Lernperiode ist das Routing nicht optimal.
 - Bei häufigem Neubeginn nehmen viele Pakete Wege unbekannter Qualität.
 - Bei seltenem Neubeginn ergibt sich ein schlechtes Anpassungsverfahren.



Routing in der Gruppenkommunikation (Multicast)

- Jedes Paket enthält entweder
 - 1. eine Liste der Bestimmungsorte, oder
 - 2. eine Bitleiste, welche die Bestimmungsorte angibt, oder
 - 3. alle Empfänger werden anhand einer Kennung (Gruppenadresse) identifiziert (Bsp.: IP-Multicast)
- Jeder Knoten bestimmt aus den in einem Paket enthaltenen Bestimmungsorten die Menge der Ausgabeleitungen.
- Für jede zu benutzende Ausgabeleitung wird eine Kopie des Pakets erzeugt.
 - Diese Kopie enthält nur die Bestimmungsorte, die über diese Leitung erreicht werden sollen (bei Fall 1 bzw. 2)

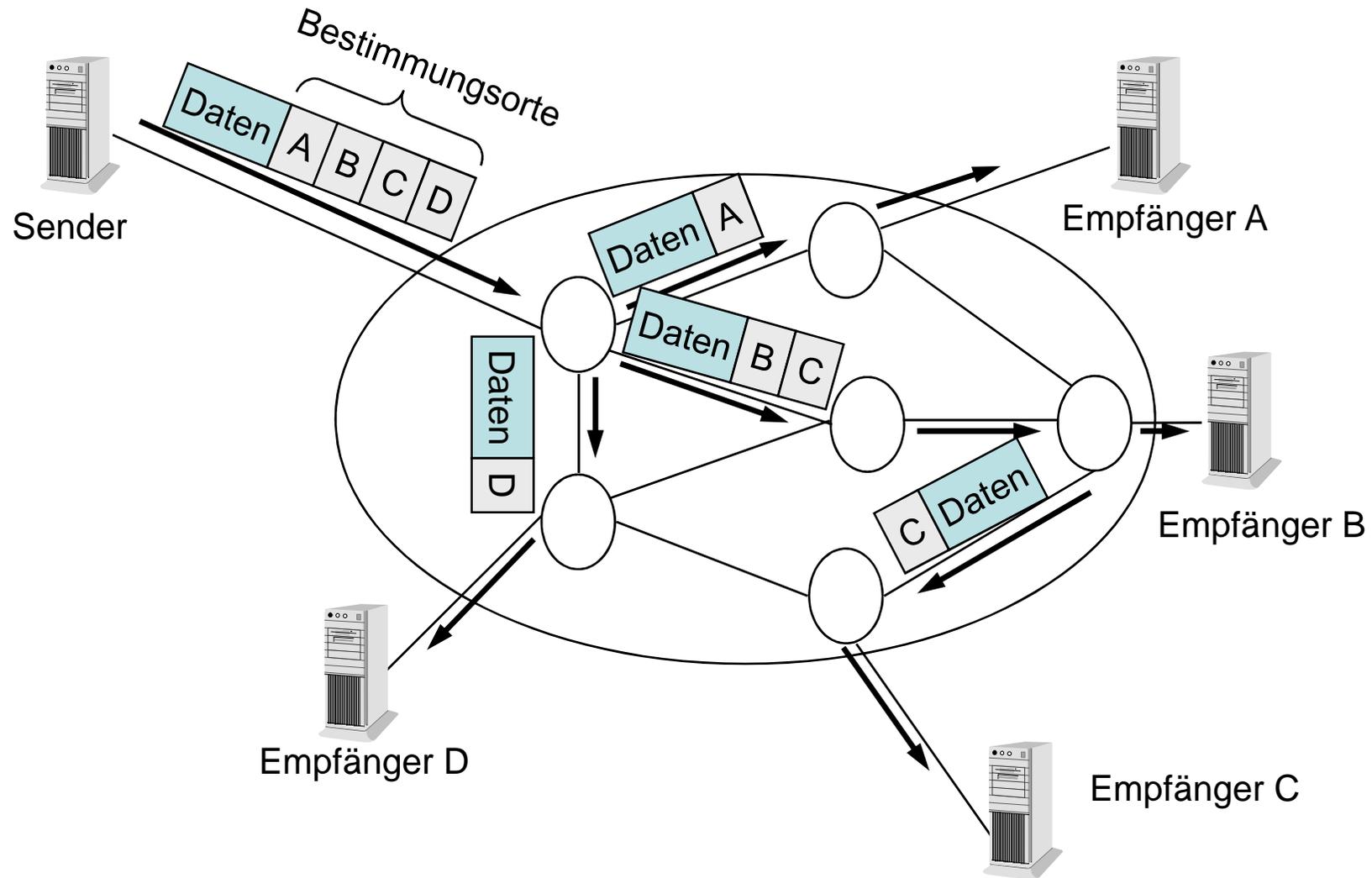


Routing in der Gruppenkommunikation

- Die Pakete müssen an mehreren Ausgängen ausgegeben werden
- Problem der Gruppenadressierung:
 - Ansatz 1 - Explizite Adressierung:
Jeder Empfänger wird im Paket explizit angegeben
 - Liste der Bestimmungsorte, bzw.
 - Bitleiste, welche die Bestimmungsorte angibt
 - Dem Sender (und dem Router) müssen alle Empfänger bekannt sein
 - Weg zu den Empfängern wird durch Unicast-Routing-Tabelle ermittelt
 - Ansatz 2 – Gruppenadressierung:
Alle Empfänger werden anhand einer Kennung (Gruppenadresse) identifiziert (Bsp.: IP-Multicast)
 - Dem Sender (und dem Router) sind die Empfänger i.Allg. nicht bekannt
 - Einsatz eines separaten Routing-Protokolls (eigene Multicast-Routing-Tabelle)



Multicast-Routing mit Empfängerliste





Multicast-Routing: Reverse-Path-Forwarding

□ Algorithmus

- 'Routing(i)' bestimmt die Ausgabeleitung für Paket mit Ziel i
- 'Incoming Link' sei die Ankunftsleitung eines Pakets
- 'Source Node' sei der Ursprungsknoten eines Pakets

```
if Destination Node = {All Nodes or Group Address}
then   if Incoming Link = Routing(Source Node)
/*Paket kommt auf dem vermutlich kürzesten Weg vom Source Node*/
   then   Outgoing Link Set := All Links - {Incoming Link} // Weiterleiten
   else   Outgoing Link set :=  $\emptyset$  // Verwerfen des Pakets
else   Outgoing Link := Routing(Destination Node)
```



Verteiltes adaptives Routing

- Jeder Knoten tauscht periodisch Routing-Informationen mit jedem seiner Nachbarn aus.

- Typischerweise unterhält jeder Knoten eine Routing-Tabelle, die für jeden anderen Knoten im Netz, bzw. für jedes Teilnetz (oder Zusammenfassungen von Teilnetzen) einen Eintrag enthält.
Hierzu zählen
 - Bevorzugte Übertragungsleitung für diesen Knoten / dieses Teilnetz
 - Schätzung über Zeit oder Entfernung zu diesem Knoten / diesem Teilnetz:
 - Anzahl hops,
 - Geschätzte Verzögerung in Millisekunden sowie
 - Geschätzte totale Anzahl von Paketen, die entlang des Weges warten.



Verteiltes adaptives Routing: Varianten

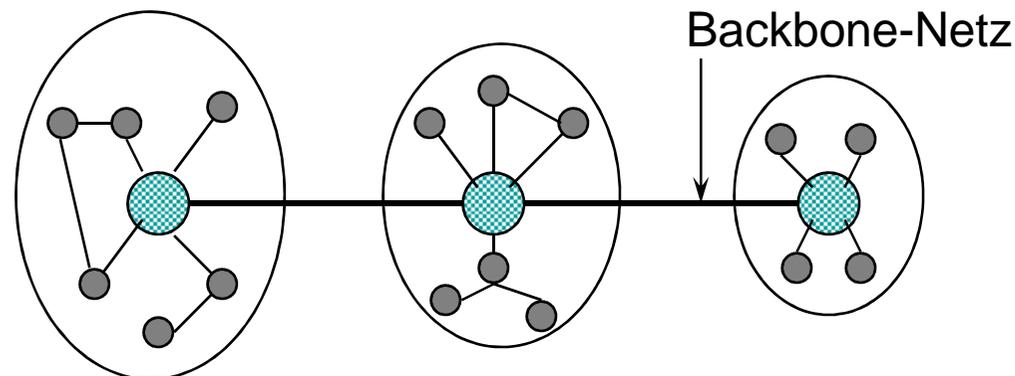
- Die Schätzungen werden gewonnen aus
 - Zeit oder Entfernung zu den Nachbarn (z.B. aus speziellen Echopaketen mit Zeitstempeln)
 - Schätzungen der Nachbarn

- Varianten
 - Synchroner Austausch von Routing-Informationen in bestimmten Aktualisierungsintervallen
 - Asynchroner Austausch bei signifikanten Änderungen



Hierarchisches Routing

- Grundlage
 - Aufteilung großer Netze in Regionen
 - Die Knoten in einer Region haben nur Routing-Informationen über ihre eigene Region.
 - In jeder Region gibt es zumindest einen ausgezeichneten Knoten, der als Schnittstelle zu anderen Regionen dient.
- In sehr großen Netzen sind weitere Hierarchieebenen möglich
 - Regionen, Cluster, Zonen, Gruppen, ...



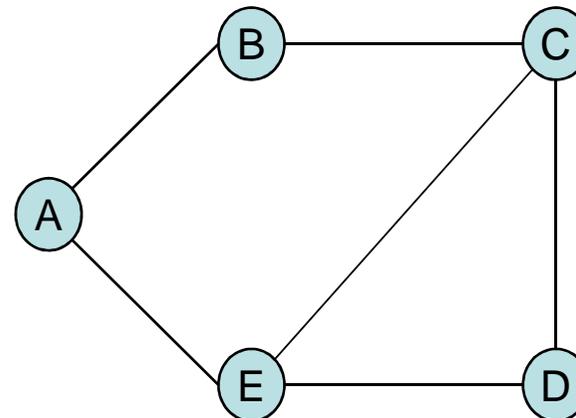


Distance Vector Routing

- **Verteiltes, adaptives Routing**
- Als RIP (Routing Information Protocol) im Internet (in kleinen Netzen) eingesetzt
- Jeder Router speichert eine Tabelle mit der besten Entfernung (z.B. Anzahl Hops, Verzögerung in ms) zu jedem Ziel und dem dazugehörigen Ausgang bzw. nächstem Hop
- Benachbarte Router teilen sich in regelmäßigen Abständen den Inhalt ihrer Routing-Tabelle mit und aktualisieren damit ihre eigene Tabelle (Bellman-Ford-Algorithmus)

Routing-Tabelle von A:

| Ziel | Nächster Hop | Entfernung |
|------|-------------------|------------|
| B | - | 1 |
| C | B/E ¹⁾ | 2 |
| D | E | 2 |
| E | - | 1 |

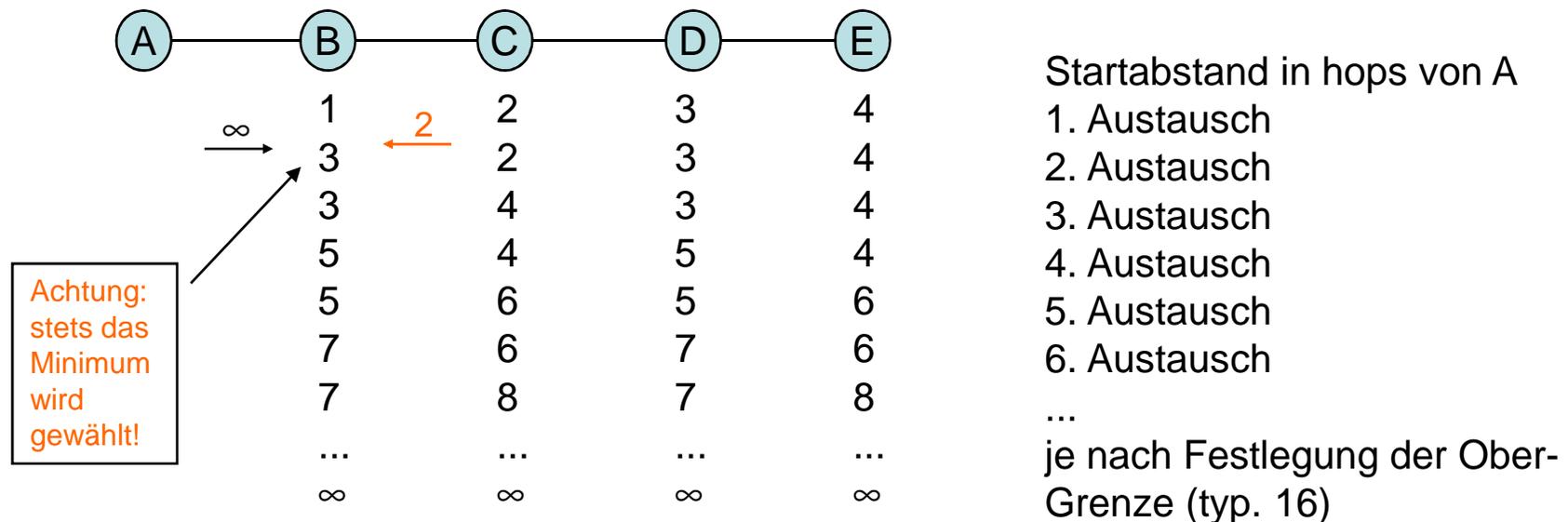


¹⁾ je nach dem, welche Route zuerst bekannt ist



Distance Vector Routing (Fortsetzung)

- Problem bei Distance Vector Routing:
 - Für Netze mit vielen Routern: langsame Konvergenz zu einem konsistenten Zustand wegen „count-to-infinity“ –Problematik
- Beispielszenario: Router A bis E verbunden, plötzlich fällt A aus



- Diverse Lösungsansätze:
 - Split Horizon, Poisoned Reverse, Triggered Updates, Path Vector



Distance Vector Routing

Lösungsansätze für schnellere Konvergenz

- Split Horizon
 - Eine Pfadinformation darf nicht über das selbe Interface veröffentlicht werden, worüber sie empfangen wurde (simple split horizon).
- Poisoned Reverse
 - Die Pfadinformation wird zwar an das Interface zurückgeschickt, über die sie empfangen wurde, aber die Entfernung wird auf unendlich gesetzt.
 - Beispiel zu voriger Folie: C darf Route nach A nicht an B weitergeben (bzw. nur mit Entfernung unendlich), weil diese Route von B annociert wurde
 - funktioniert nicht für größere Schleifen
- Triggered Updates
 - Ändert sich eine Metrik, wird die Pfadinformation sofort propagiert und nicht erst nach einem Timeout.
- Path Vector
 - Es wird nicht nur der nächste Hop, sondern der ganze Pfad bis zu Ziel weitergegeben bzw. gespeichert
 - Kommt bei BGP zum Einsatz



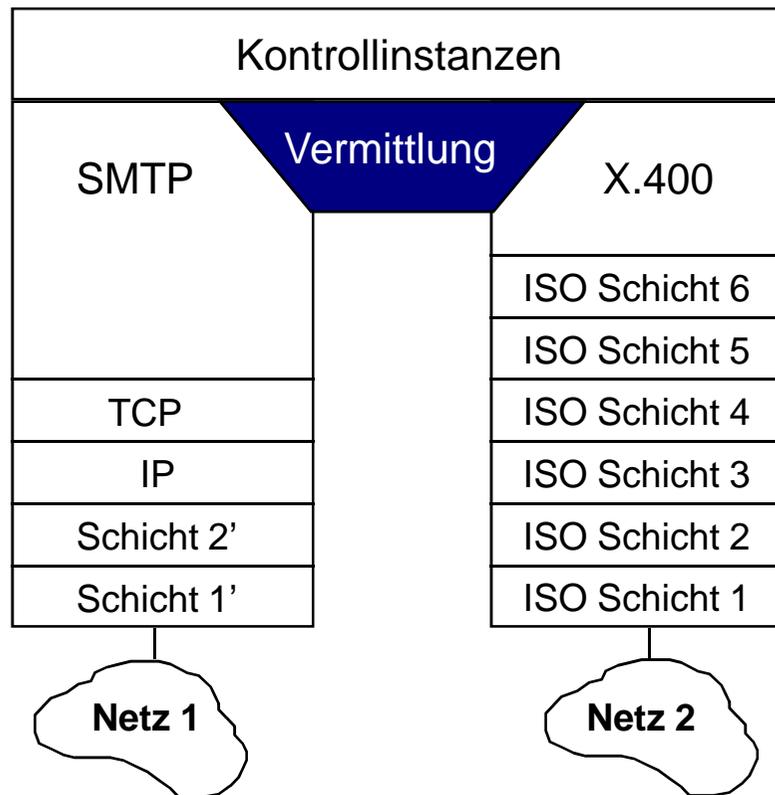
Link State Routing

- **Verteiltes, adaptives Routing**
- Als OSPF (Open Shortest Path First) und IS-IS (Intermediate System - Intermediate System) im Internet eingesetzt
- Algorithmus:
 - Entdecken neuer Nachbarn über HELLO-Pakete
 - Bestimmung der Link-Kosten:
 - entweder durch Konfiguration vorgegeben, oder ermittelt über Messung der Verzögerung zu jedem Nachbarn (ECHO-Paket misst Umlaufzeit)
 - Erstellen eines „Link-State“-Paketes mit allen gelernten Daten
 - beinhaltet ID des Senders (Routers), Liste der Nachbarn mit Verzögerung, Alter
 - periodische oder ereignisgesteuerte (z.B. neuer Nachbar, Ausfall) Erzeugung
 - Aussenden dieses Paketes an alle Nachbarn
 - Nachbarn geben Link-States wiederum an ihre Nachbarn weiter usw.
→ prinzipiell Fluten an alle Router, aber mit Verfeinerungen: Vernichten von Duplikaten, Zerstören der Information nach gewissem Alter etc.
 - Jeder Router kennt am Ende die Topologie des Netzwerkes
 - Berechnung des kürzesten Pfades zu allen anderen Routern (z.B. Dijkstra)
 - sehr rechenaufwendig, Optimierungen existieren



Gateways

- Kopplung von Netzwerken auf einer höheren Schicht (≥ 4)
- falls Quell- und Zielrechner verschiedene höhere (d.h. für Anwendung sichtbare) Protokolle verwenden
- Typisches Beispiel: Gateway zwischen verschiedenen Mail-Systemen



SMTP: Simple Mail Transport Protocol



Literatur zu Routingverfahren (Auswahl)

- **Huitema, C.:**
Routing in the Internet
2. Auflage, Prentice Hall, Inc., New Jersey, 2000,
ISBN 0-13-022647-5
Gute Übersicht über Routingverfahren im Internet

- **Tanenbaum, A. S.:**
Computer Networks
4. Auflage, Prentice Hall, Inc., New Jersey, 2002,
ISBN 0-13-394248-1
Insbesondere Kap. 5

- **Perlman, R.:**
*Interconnections Second Edition: Bridges, Routers, Switches, and
Internetworking Protocols*
Addison-Wesley, Reading, Mass., 1999,
ISBN 0-201-63448-1