Master Course
Computer Networks
IN2097

Prof. Dr.-Ing. Georg Carle

Chair for Network Architectures and Services
Department of Computer Science
Technische Universität München
http://www.net.in.tum.de

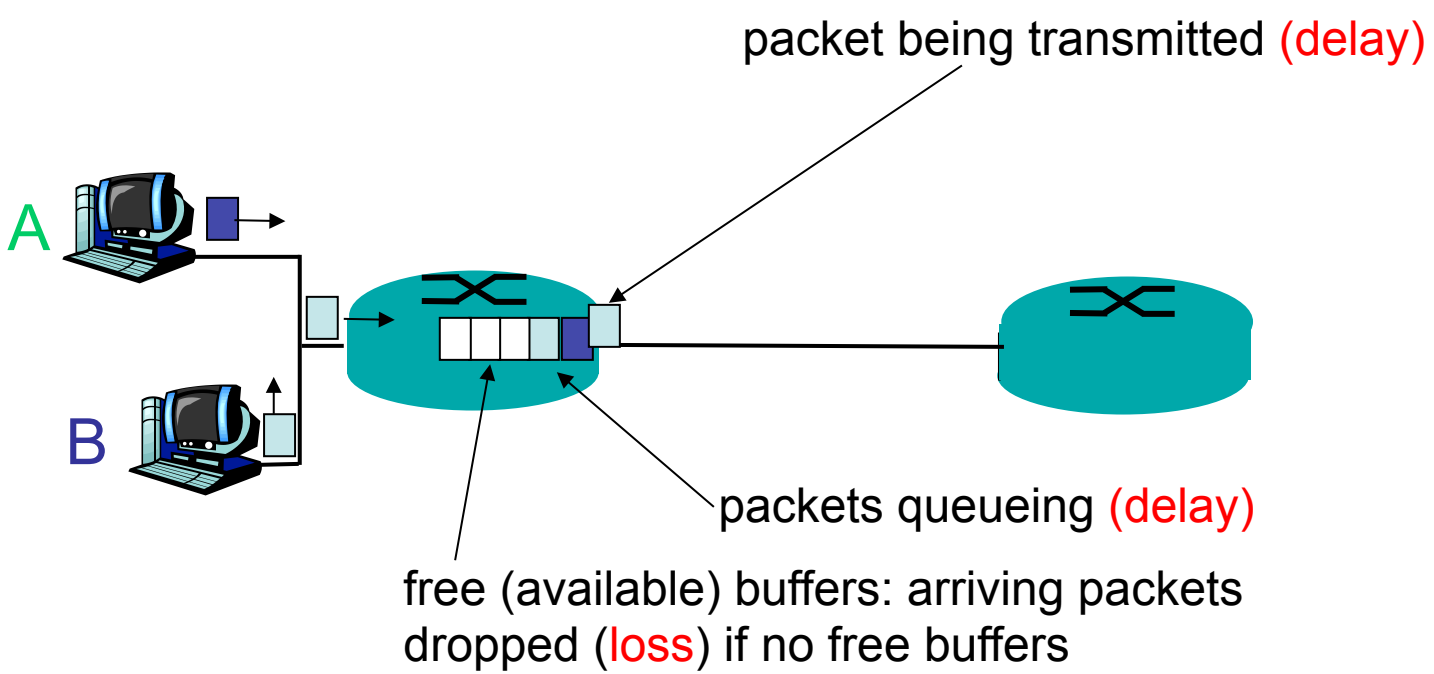
Chair for Network Architectures and Services – Prof. Carle
Department of Computer Science
TU München

Node Forwarding Performance

Reasons for Delay and Loss

Packets *queue* in router / switch buffers

- Packet arrival rate to link exceeds output link capacity
- Packets queue, wait for turn



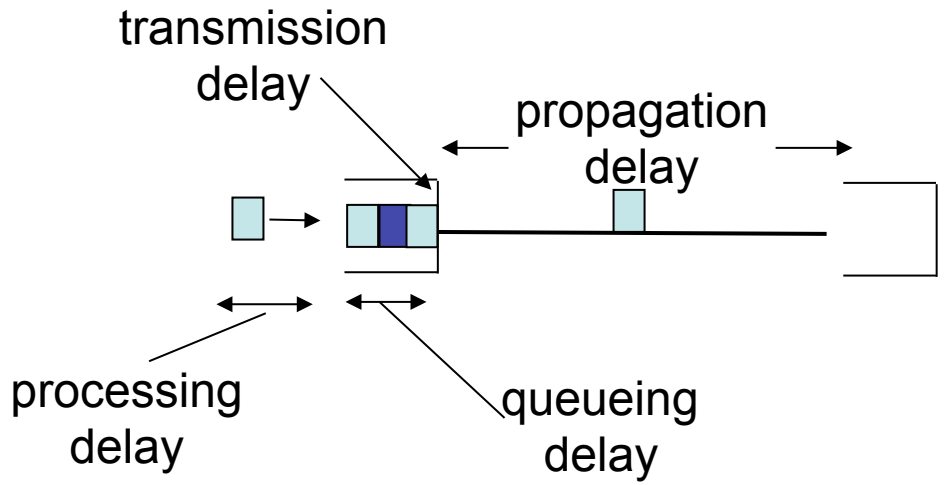
Background: Sources of Packet Delay

1. Processing delay:
   - Sending: prepare data for being transmitted
   - Receiving: interrupt handling

2. Queueing delay
   - time waiting at output link for transmission

3. Transmission delay:
   - L= packet length (bits)
   - R= link data rate (bits/sec)
   - time to send bits into link = L/R

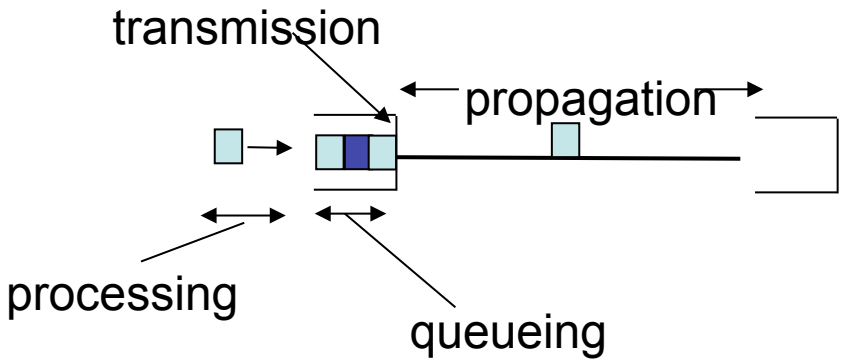
4. Propagation delay:
   - d = physical link distance
   - s = propagation speed in medium (~$2x10^8$ m/sec)
   - propagation delay = d/s



Nodal Delay

- d_{proc} = processing delay
  - typically small - a few microseconds (µs) or less
- d_{queue} = queuing delay
  - depends on congestion - may be large
- d_{trans} = transmission delay
  - = L/R, significant for low-speed links
- d_{prop} = propagation delay
  - a few microseconds to hundreds of msecs

$$d_{nodal} = d_{proc} + d_{queue} + d_{trans} + d_{prop}$$



Impact Analysis: Advances in Network Technology

data rate	trans. delay (1bit)	phys. length (1bit)	trans. delay (1kbyte)	phys. length (1kbyte)
1 Mbit/s	1 us	200 m	8 ms	1600 km
10 Mbit/s	100 ns	20 m	0,8 ms	160 km
100 Mbit/s	10 ns	2 m	80 us	16 km
1 Gbit/s	1 ns	0,2 m	8 us	1600 m
10 Gbit/s	100 ps	0,02 m	0,8 us	160 m
100 Gbit/s	10 ps	0,002 m	80 ns	16 m

- Assessment
  - Transmission delay becomes less important
    ⇨ over time; in the core of the network
  - Distance becomes more important
    ⇨ matters for communication beyond data center
  - Network adapter latency less important
    ⇨ latency of communication software becomes important

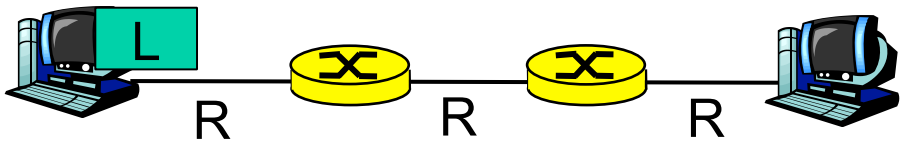
Propagation Delay

- Propagation speed: $2x10^8$ m/sec
- Transmission of packet with 625 byte (= 5000 bit):
  t= L/R=5000 / 1Gbit/s = 5 us

distance	propagation delay	equivalent transmission delay (625 byte)	CPU cycles per packet (1 GHz)	CPU cycles per byte (1 GHz)
100 m	500 ns	10 Gbit/s	500	<1
1 km	5 us	1 Gbit/s	5.000	8
10 km	50 us	100 Mbit/s	50.000	80
100 km	500 us	10 Mbit/s		800
1.000 km	5 ms	1 Mbit/s		8.000
10.000 km	50 ms	100 Kbit/s		80.000

- Suggestion for home exercise: plot graphs

Store-and-Forward vs. Circuit Switching



- Transmission delay:
  L=packet length (bits)
  R=link bandwidth (bps)
  time to transmit packet of L bits on to link with R bps = L/R
- Store and forward: entire packet must arrive at router before it can be transmitted on next link:
- Total transmission delay = 3L/R

Example: Large Message L

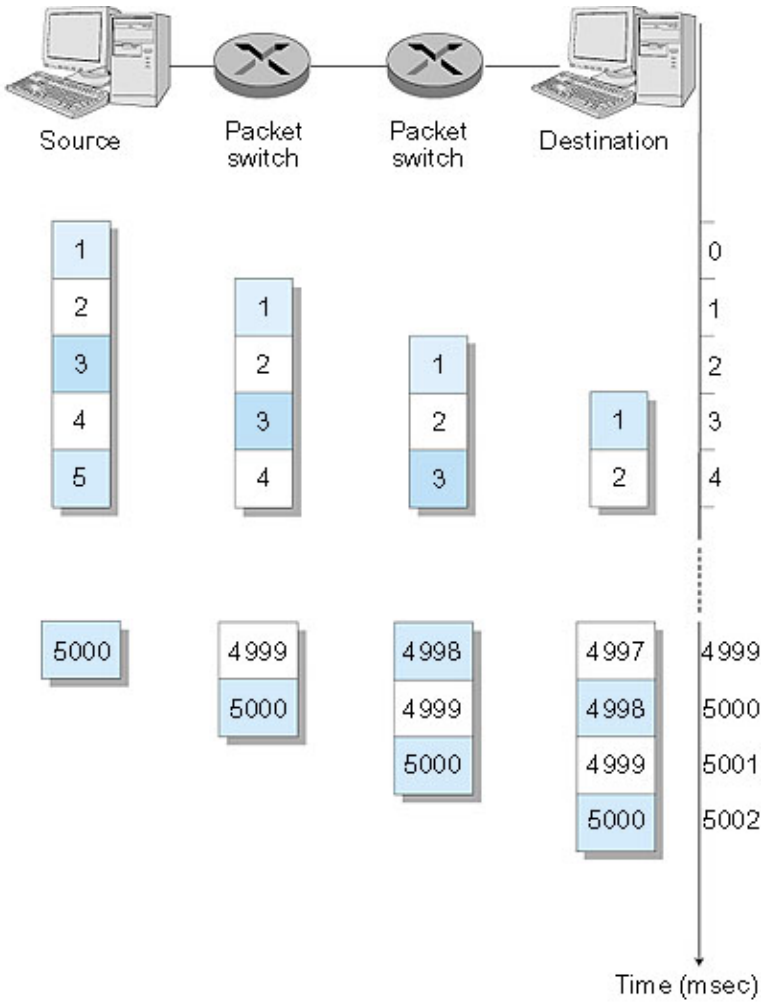
Store-and-Forward:

- L = 7.5 Mbit
- R = 1.5 Mbit/s
- Transmission delay = 15 s

Circuit Switching:

- L = 7.5 Mbit
- R = 1.5 Mbit/s
- Transmission delay = 5 s

Packet Switching: Message Segmenting



Now break up the message into 5000 packets

- Each packet 1,500 bits
- 1 msec to transmit packet on one link
- *Pipelining:* each link works in parallel
- Delay reduced from 15 sec to 5.002 sec (as good as circuit switched)
- Advantages over circuit switching?
- Drawbacks (of packet vs. Message)

Discussion

- What is the role of header lengths?
- What is the role of header compression?
- What is the cost of tunneling?
- What are the benefits of overprovisioning?
- Can you „imagine“ a visualisation of packets being transmitted over different types of links?

Questions

- Why/when is circuit switching expensive?
- Why/when is packet switching cheap?
- Is best effort packet switching suitable to carry voice communication?
- What happens if we introduce “better than best effort” service?
- How can we charge fairly for Internet services: by time, by volume, or flat?

Chair for Network Architectures and Services – Prof. Carle
Department of Computer Science
TU München

Connection-Oriented Network Architectures

Outline: Connection-Oriented Network Architectures

- Principles of connection oriented networks
- Representative connection-oriented technology: ATM
- Virtual Private Networks (VPNs)

Connection-Oriented Networks - Connection Setup

- In addition to routing and forwarding, *connection-setup* is 3rd important function in some network architectures:
  - X.25, Frame Relay, ATM, MPLS, GMPLS
- Before datagrams flow, two end hosts and intervening switches establish virtual connection
  - Switches get involved in connection establishment
- Network layer vs. transport layer connection-oriented service:
  - Network layer: connection between hosts, or routers
  - Transport layer: connection between two processes

Connection-orientend vs. Connection-less Network Service

- Datagram network provides connection-less network-layer service (example: Internet)
- Virtual Circuit network provides connection-oriented network-layer service (example: MPLS network)
- Analogous to the transport-layer services, but:
  - service: host-to-host, or edge-node-to-edge-node
  - no choice: typically, network provides one or the other
  - implementation: typically in the network core

Virtual Circuits

"source-to-destination path behaves much like telephone circuit"

- network actions along source-to-destination path
- performance-wise *(this is not a necessary property)*

- Two-stage process
  - Connection setup *before* data can flow to establish "connection state" in switches between source and destination hosts
  - Data transfer
- Each packet carries VC Identifier (VCI), not destination host address
- *Every* switch on source-to-destination path maintains "state" for each passing connection
- Link, switch resources (bandwidth, buffers) may be *allocated* to VC (dedicated resources = predictable service)

Virtual Circuits

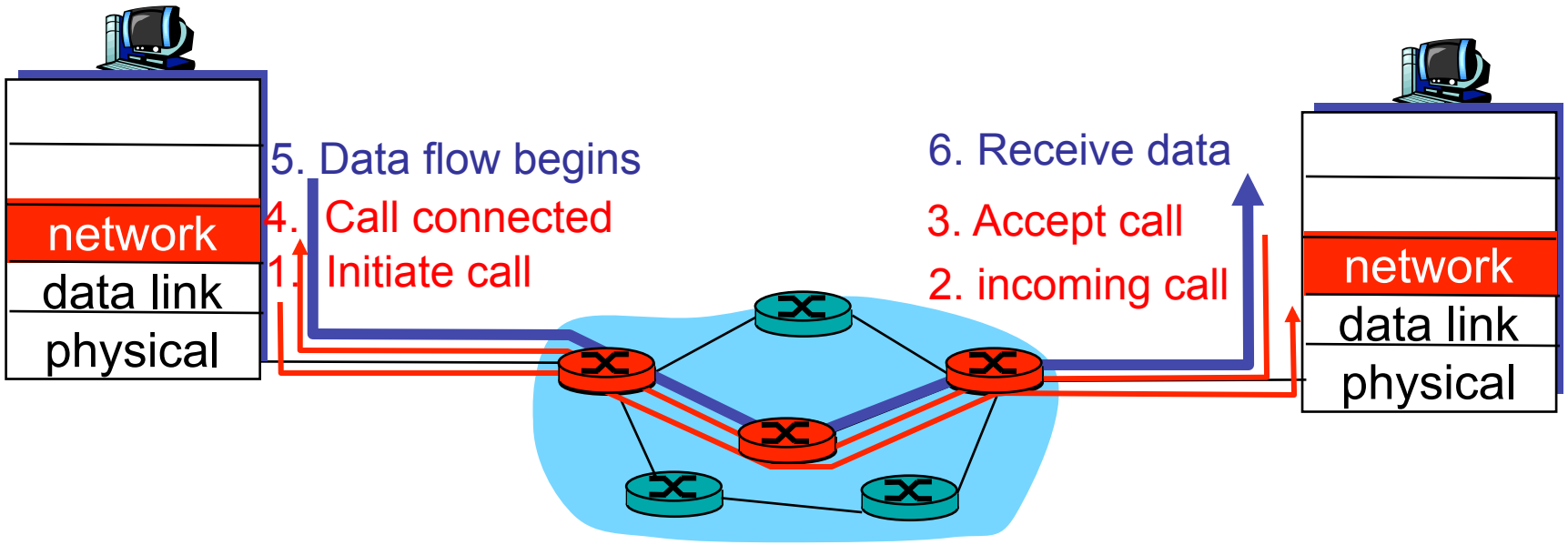
Alternative approaches to establish connection state

- Network administrator configures state
  - Virtual circuit is permanent (PVC)
  - Network administrator can delete PVC
  - Is long-lived or administratively configured VC
- A host can send messages into the network to cause the state to be established
  - This is referred as signalling and the resulting virtual circuit is said to be switched (SVC)
  - A host may set up and delete such a VC dynamically without involvement of a network administrator

Virtual Circuits: Signaling Protocols

Signaling

- Used to setup, maintain, and teardown VC
- Used in X.25, Frame-Relay, ATM, MPLS, GMPLS



Network Service Models

Q: What *service model* for "channel" (Virtual Circuit) transporting datagrams from sender to receiver?

- Virtual circuits can have specific quality of service (QoS)
- Network can give guarantee ⇨ switches reserve resources

Example services for individual datagrams:

- guaranteed delivery
- guaranteed delivery with less than 40 msec delay

Example services for a flow of datagrams:

- in-order datagram delivery
- guaranteed minimum bandwidth to flow
- restrictions on changes in inter-packet spacing

Virtual Circuit Implementation

A VC consists of:

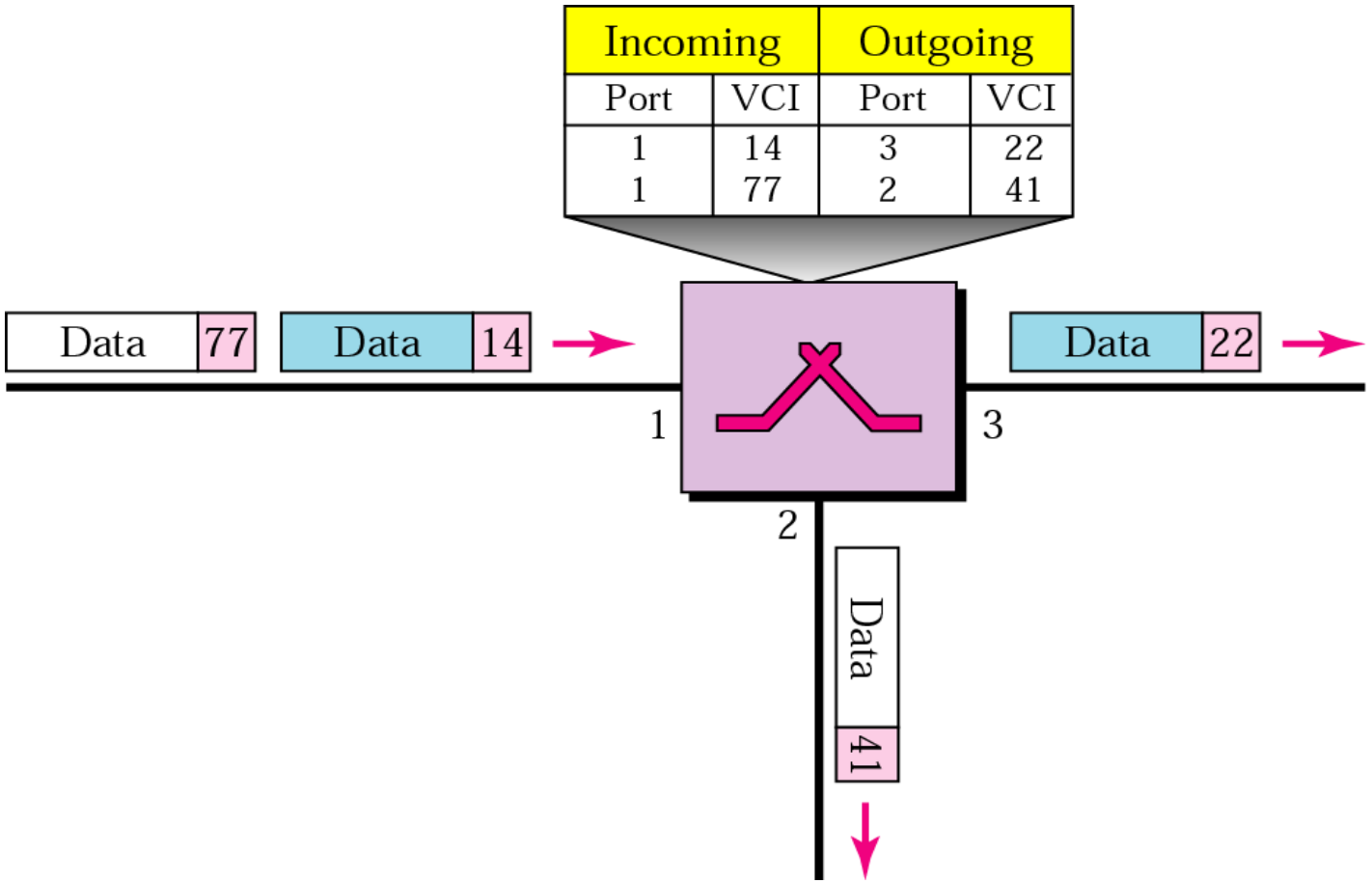
1. path from source to destination
2. "labels": VC numbers (virtual circuit identifiers – VCIs), one number for each link along path
3. entries in forwarding tables in switches along path

- Packet belonging to VC carries VC number (rather than destination address)
  - destination address is used in connection setup message of signaling protocol)
- VC number can be changed on each link – "label swapping"
  - If VCI is changed, new VCI comes from forwarding table
  - VCI is not a globally unique identifier for the connection; rather it has significance only on a given link

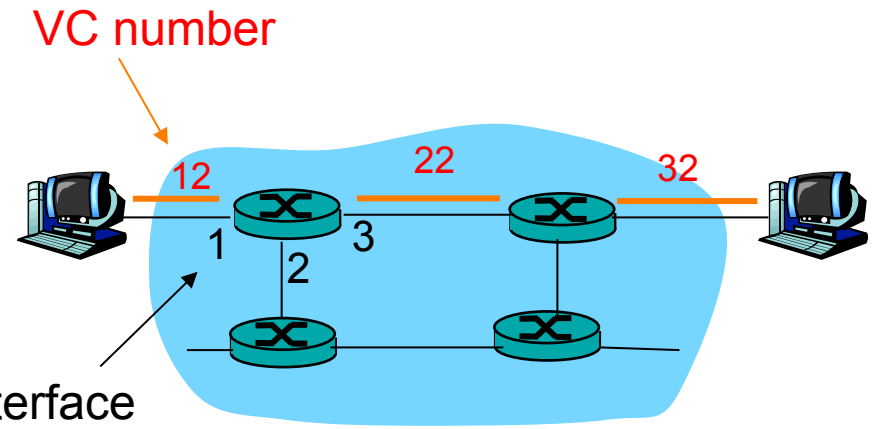
Virtual Circuit Implementation

- The combination of
  VCI of the packets received at the switch
  and the interface on which they are received
  uniquely identifies the virtual connection
- When a new connection is created,
  a new VCI for that connection must be assigned
  on each link of the connection
  - The chosen VCI on a given link must not be in use
    on that link by some existing connection

Virtual Circuit Switching



VC Network Forwarding Table



Forwarding table in northwest VC node:

Incoming port	Incoming VC #	Outgoing port	Outgoing VC #
1	12	3	22
2	63	1	18
3	7	2	17
1	97	3	87
…	…	…	…

VC node maintain connection state information!

ATM – Asynchronous Transfer Mode

- Connection-oriented packet-switched network
- Packets are called cells
  - 5 byte header + 48 byte payload
- Fixed length packets are easier to switch in hardware
  - Simpler to design
  - Enables parallelism
- Short packets have low transmission delay
  - No queuing of short packets behind long packets currently being transmitted
  - Low per-switch delays possible

ATM Service Models

ATM Services: transport cells across ATM network

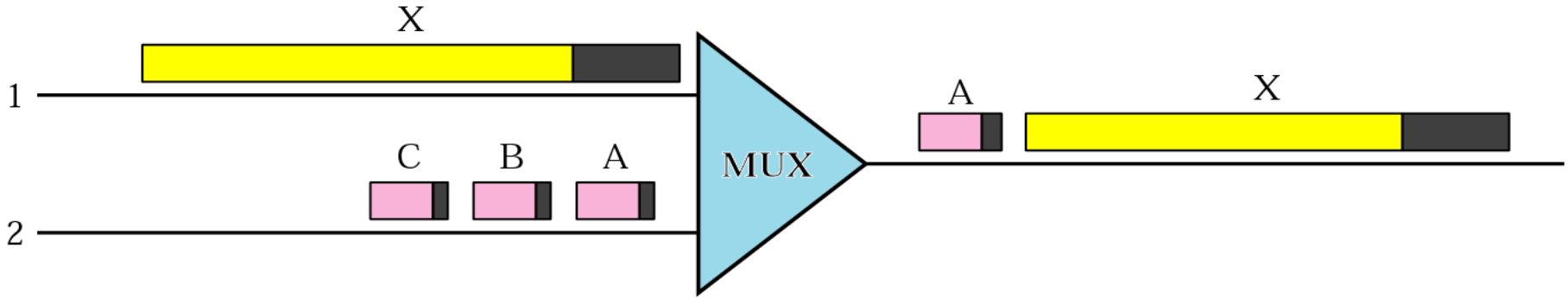
- very different services than IP network layer
- possible Quality of Service (QoS) Guarantees

Network Architecture	Service Model	Guarantee				Congestion feedback
		Bandwidth	Loss	Order	Timing	
Internet	best effort	none	no	no	no	no (inferred via loss)
ATM	CBR	constant rate	yes	yes	yes	no congestion
ATM	VBR	guaranteed rate	yes	yes	yes	no congestion
ATM	ABR	guaranteed minimum	no	yes	no	yes
ATM	UBR	none	no	yes	no	no

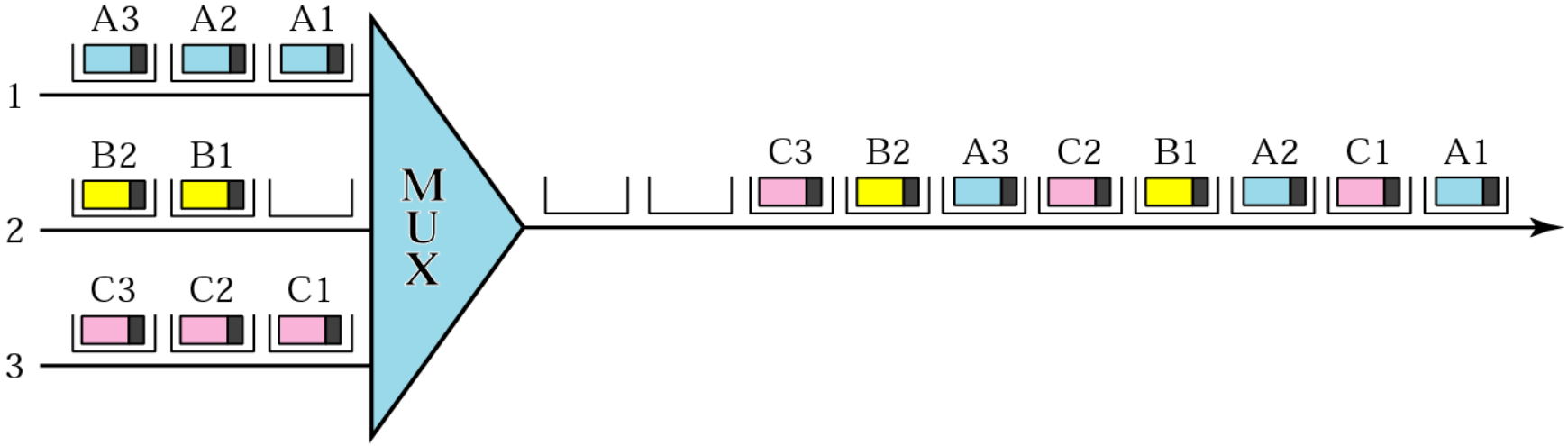
CBR: Constant Bit Rate
VBR: Variable Bit Rate
ABR: Arbitrary Bit Rate
UBR: Unspecified Bit Rate

Multiplexing of Variable vs. Fixed Size Packets

- Multiplexing of variable size packets

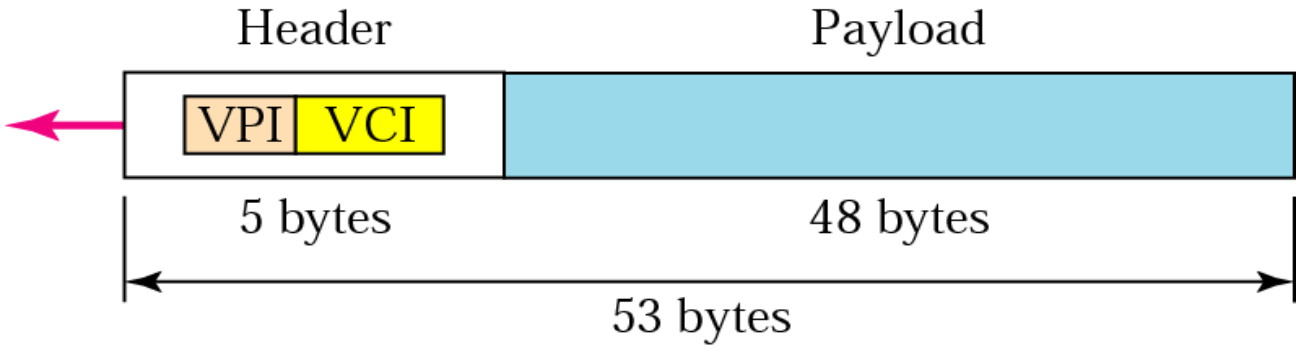


- ATM Multiplexing

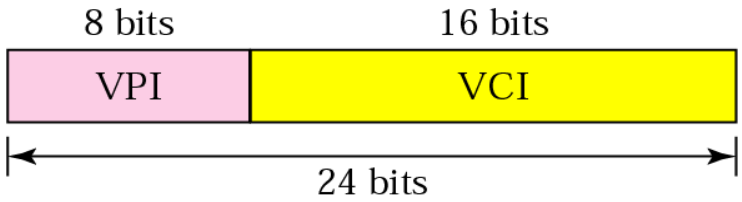


ATM Identifiers

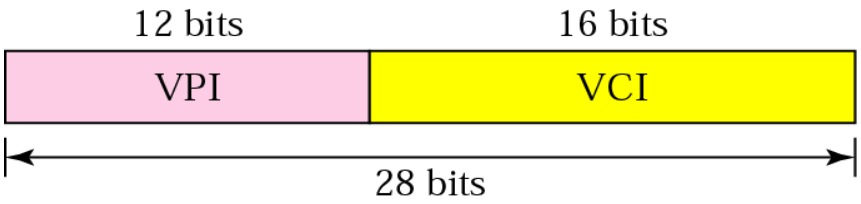
- ATM Cell



- Virtual Path Identifiers and Virtual Channel Identifiers



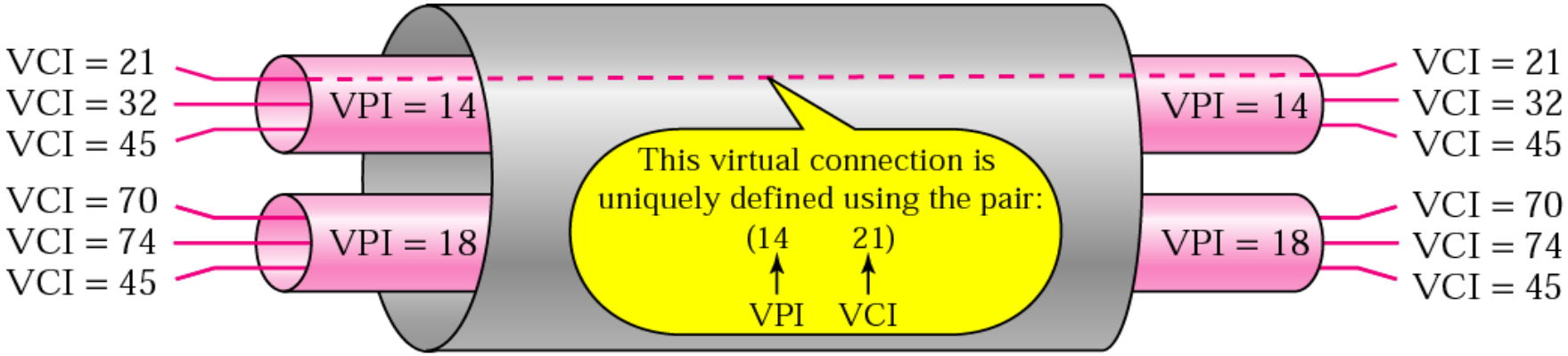
a. VPI and VCI in a UNI



b. VPI and VCI in an NNI

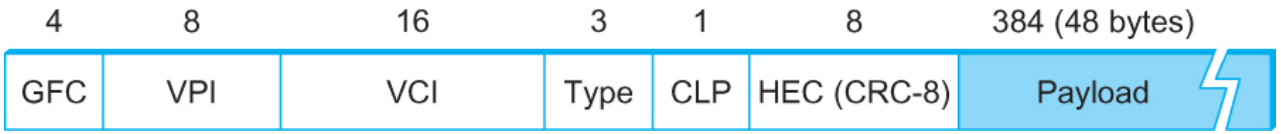
(UNI: User-to-Network-Interface
NNI: Network-to-Network-Interface)

ATM Virtual Connections



ATM Cell Format

- ATM
  - User-Network Interface (UNI)
    - Host-to-switch format
    - GFC: Generic Flow Control
    - VCI: Virtual Circuit Identifier
    - Type: management, congestion control
    - CLP: Cell Loss Priority
    - HEC: Header Error Check (CRC-8)



  - Network-Network Interface (NNI)
    - Switch-to-switch format
    - GFC becomes part of VPI field

ATM VCs

- Advantages of ATM VC approach:
  - QoS performance guarantee for connection mapped to VC (bandwidth, delay, delay jitter)
- Drawbacks of ATM VC approach:
  - Inefficient support of datagram traffic
  - one PVC between each source/destination pair does not scale
  - SVC introduces call setup latency, processing overhead for short lived connections

ATM Physical Layer

Physical Medium Dependent (PMD) sublayer

- **SONET/SDH**: transmission frame structure (like a container carrying bits);
  - bit synchronization;
  - several speeds:
    - OC3: 155.52 Mbps
    - OC12: 622.08 Mbps
    - OC48: 2.45 Gbps
    - OC192: 9.6 Gbps
- **other physical layers also possible**

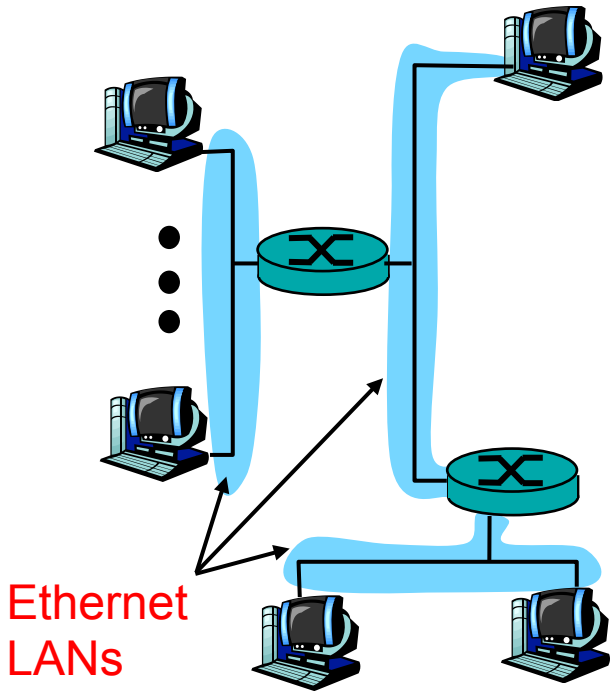
IP-Over-ATM

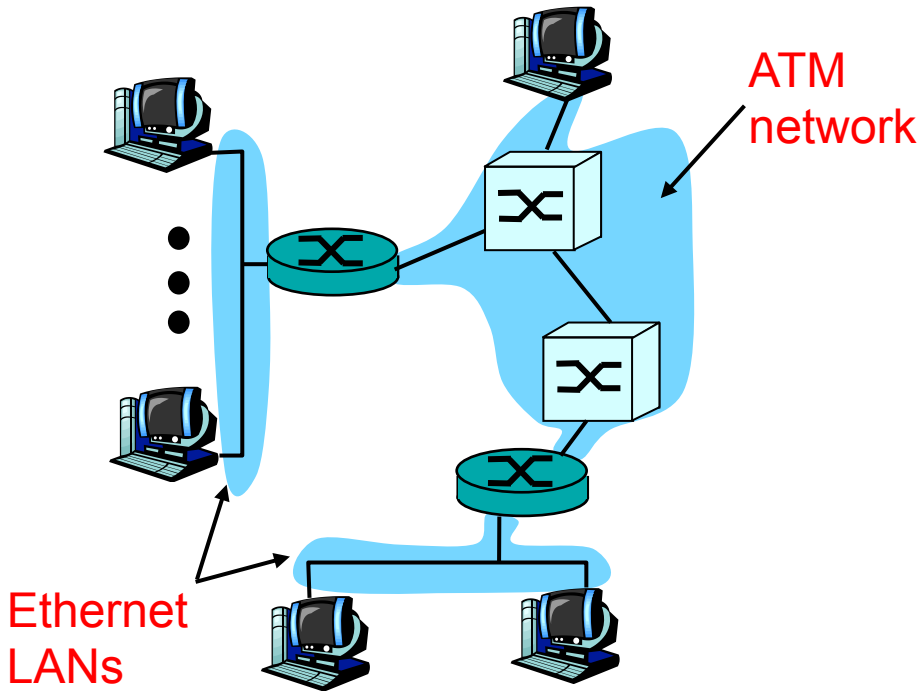
Classic IP only

- 3 “networks” (e.g., LAN segments)
- MAC (802.3) and IP addresses

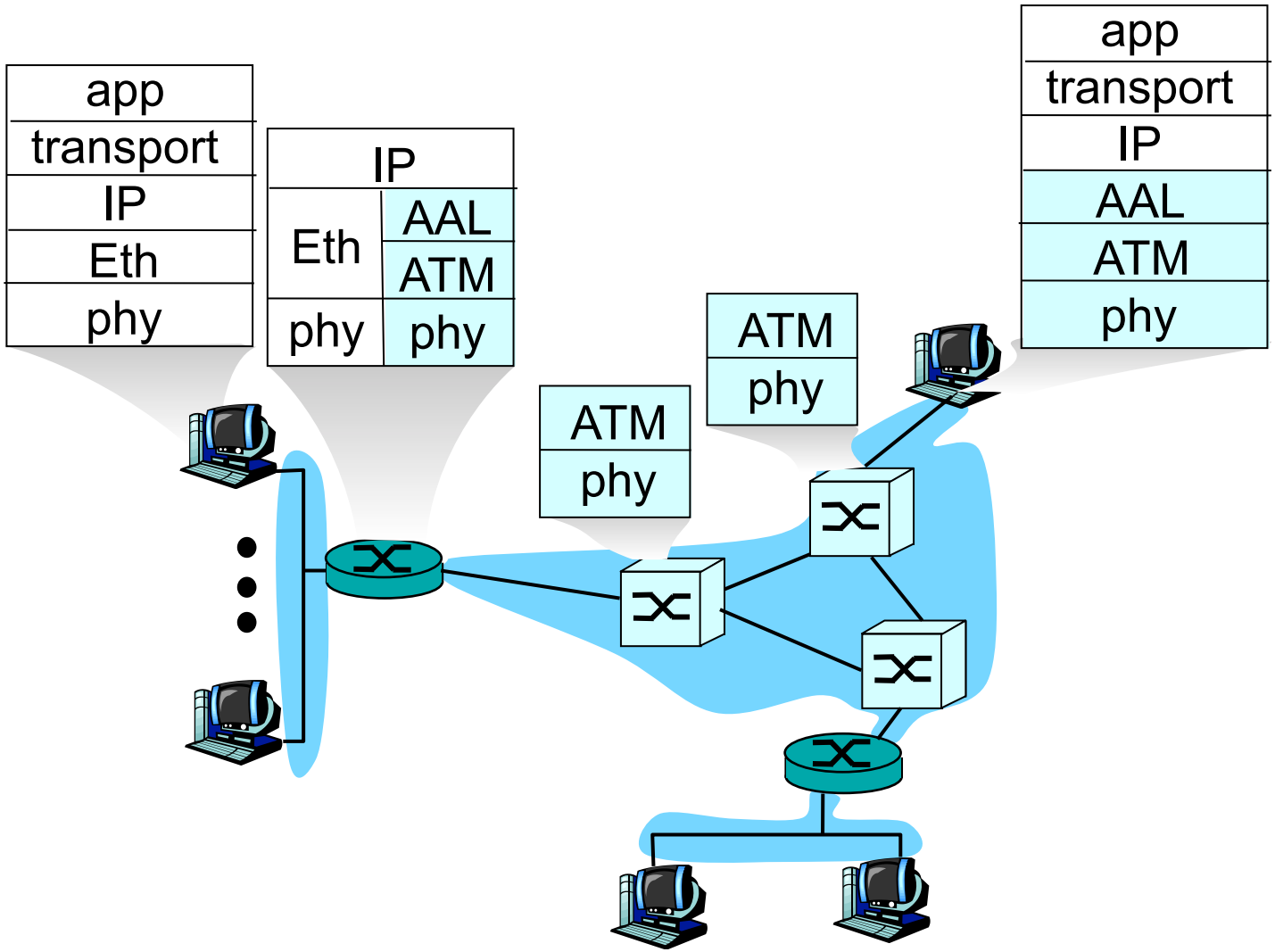
IP over ATM

- replace “network” (e.g., LAN segment) with ATM network
- ATM addresses, IP addresses





IP-Over-ATM


app
transport
IP
Eth
phy
IP
Eth
AAL
ATM
phy
phy
ATM
phy
ATM
phy
app
transport
IP
AAL
ATM
phy

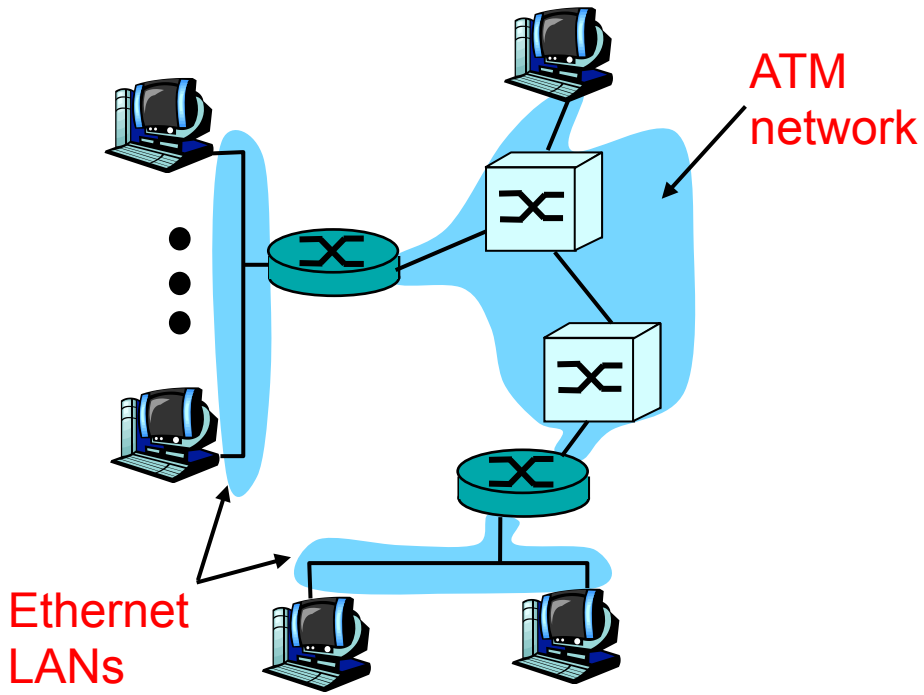
Datagram Journey in IP-over-ATM Network

- At source host:
  - IP layer maps between IP, ATM destination address (using ARP)
  - passes datagram to AAL5
  - AAL5 encapsulates data, segments cells, passes to ATM layer
- ATM network: moves cell along VC to destination
- At destination host:
  - AAL5 reassembles cells into original datagram
  - if CRC OK, datagram is passed to IP

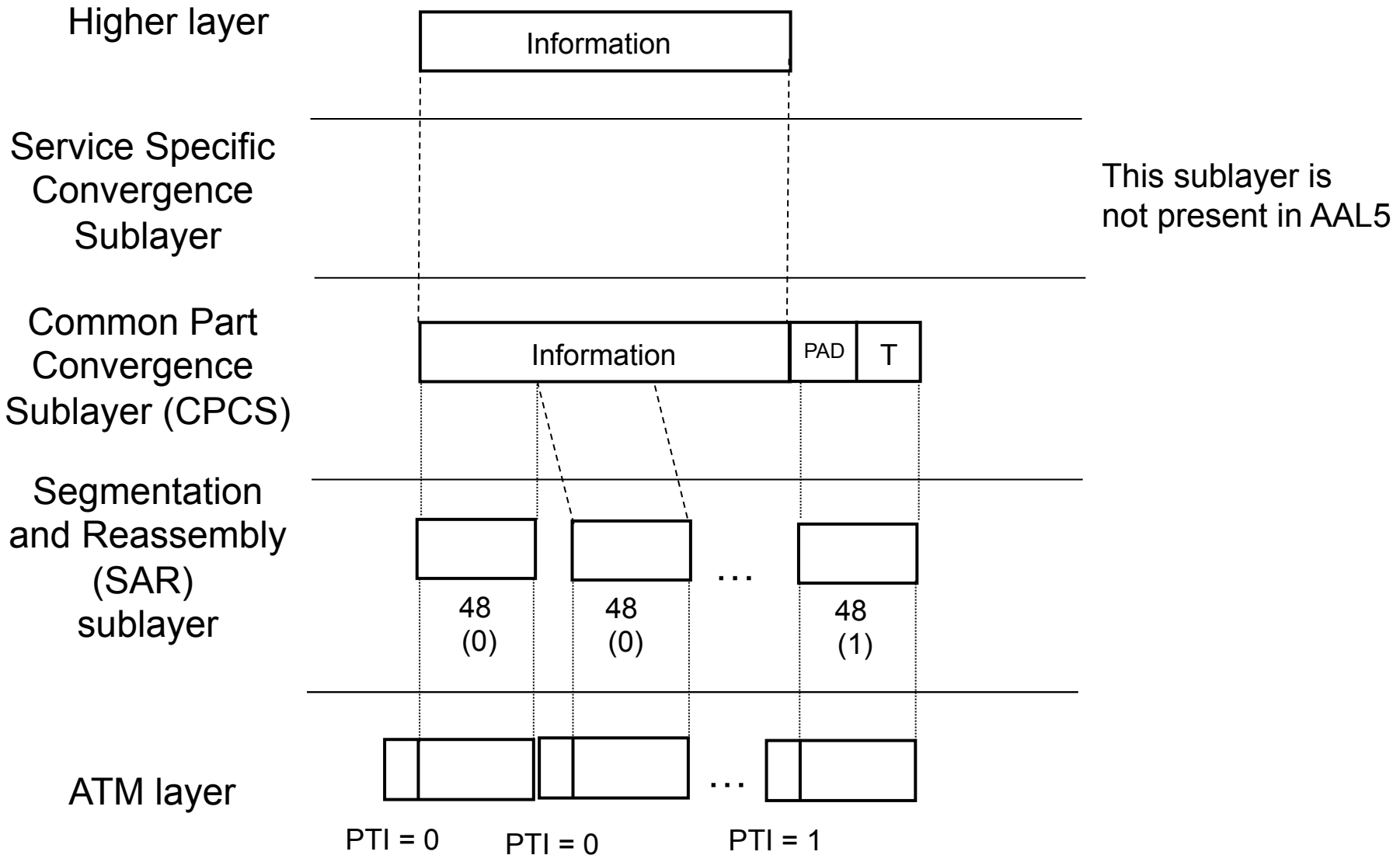
IP-Over-ATM

Issues:

- IP datagrams into ATM AAL5 PDUs
- from IP addresses to ATM addresses
  - just like IP addresses to 802.3 MAC addresses!
  - ARP server



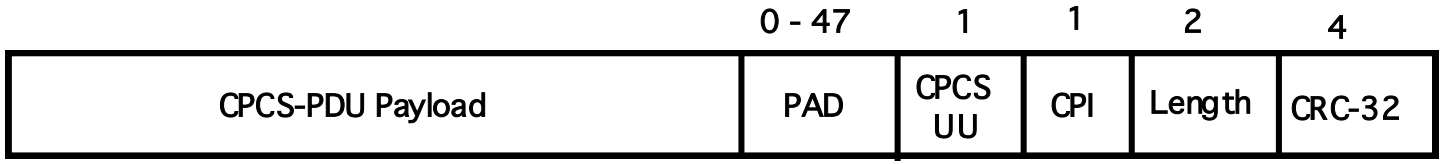
ATM Adaptation Layer: AAL 5 Layering



AAL5 Protocol

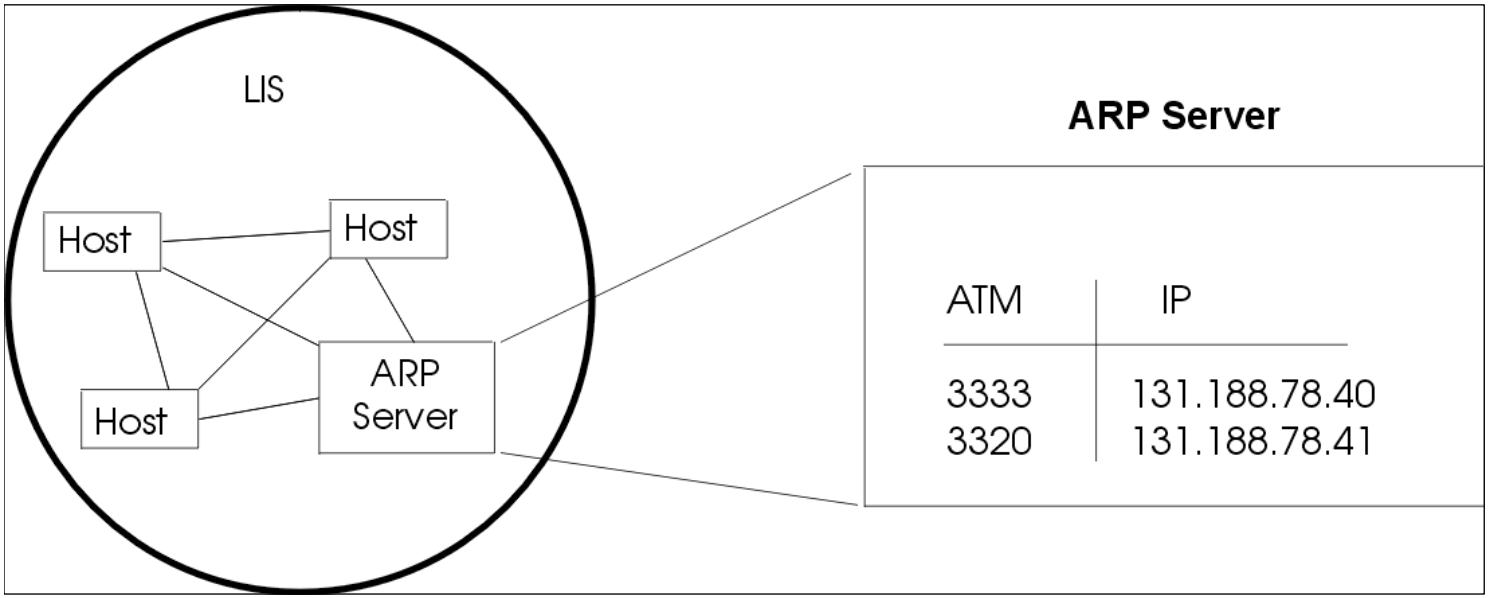
- CPCS-PDU payload length
  - can be up to 65,535 octets
  - must use PAD (0 to 47 octets) to align CPCS-PDU length to a multiple of 48 octets
- Fields in trailer *(CPCS-UU and CPI not relevant in this course)*

PAD	Padding
CPCS-UU	CPCS User-to-User Indicator
CPI	Common Part Indicator
Length	CPCS-PDU Payload Length
CRC-32	Cyclic Redundancy Chuck



Classical IP and ARP over ATM (CLIP)

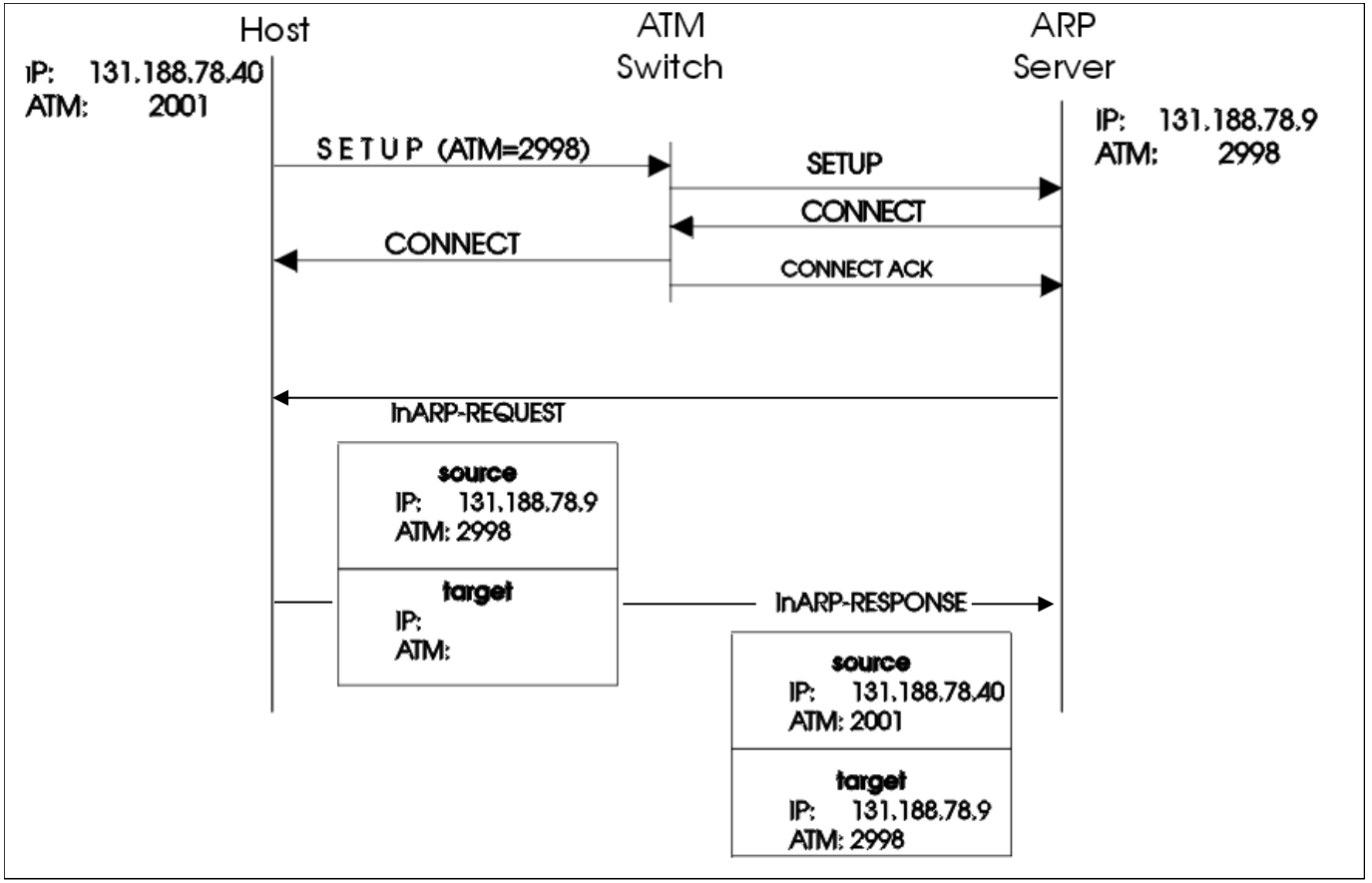
- CLIP: one of several approaches for IP services supported by ATM network
- RFC 2225: Classical IP and ARP over ATM
- Encapsulation of IP packets into AAL PDUs
- Support for large MTU sizes
- CLIP uses an ATMARP server in a Logical IP Subnet (LIS)



Classical IP and ARP over ATM (CLIP)

- RFC 2225: Classical IP and ARP over ATM
- ATMARP Server Operational Requirements
  - ATMARP server, after completion of a new VC, will transmit an InATMARP request to determine the IP address of client
  - InATMARP reply from client contains information necessary for ATMARP Server to build its ATMARP table
  - This table used to reply to ATMARP requests
- InATMARP is the same protocol as the original InARP protocol presented in RFC 1293 but applied to ATM networks: Discover the protocol address of a station associated with a virtual circuit.
- RFC 1293: T. Bradely and C. Brown, "Inverse Address Resolution Protocol“, January 1992
  - solution designed for Frame Relay and similar networks

Classical IP and ARP over ATM (CLIP)



Classical IP and ARP over ATM (CLIP)

- RFC 2225: Classical IP and ARP over ATM
- ATMARP Client Operational Requirements
  1. Initiate the VC connection to ATMARP server for transmitting and receiving ATMARP and InATMARP packets
  2. Respond to ARP_REQUEST and InARP_REQUEST packets received on any VC appropriately
  3. Generate and transmit ARP_REQUEST packets to ATMARP server and process ARP_REPLY appropriately. ARP_REPLY packets should be used to build/refresh own client ATMARP table entries.
  4. Generate and transmit InARP_REQUEST packets as needed and process InARP_REPLY packets appropriately. InARP_REPLY packets should be used to build/refresh its own client ATMARP table entries.
  5. Provide ATMARP table aging function to remove own old client ATMARP tables entries after a period of time.

Chair for Network Architectures and Services – Prof. Carle
Department of Computer Science
TU München

Virtual Private Networks

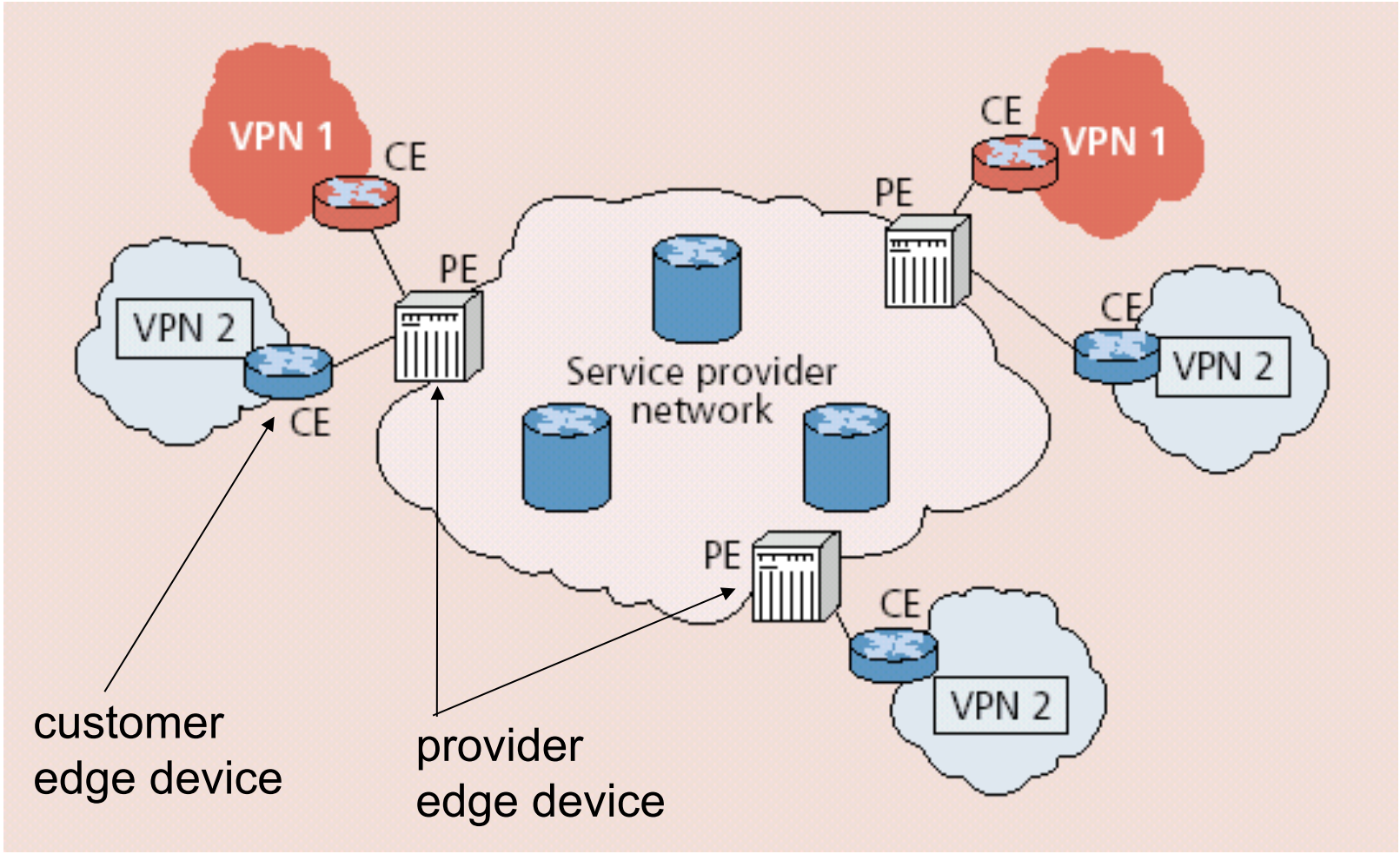
Virtual Private Networks (VPN)

VPNs

Networks perceived as being private networks by customers using them, but built over shared infrastructure owned by a service provider (SP)

- Service provider infrastructure:
  - backbone
  - provider edge devices
- Customer:
  - customer edge devices (communicating over shared backbone)

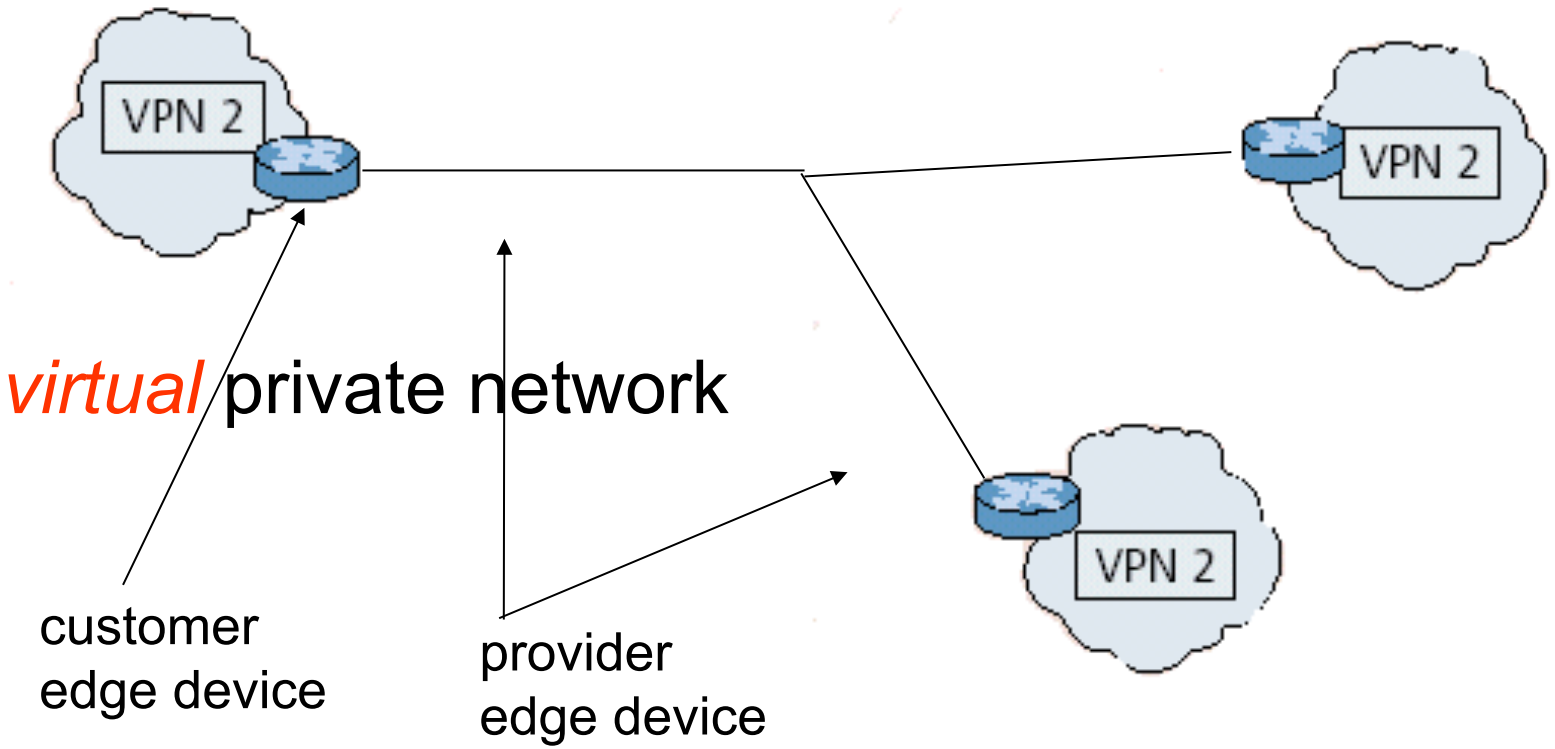
VPN Reference Architecture



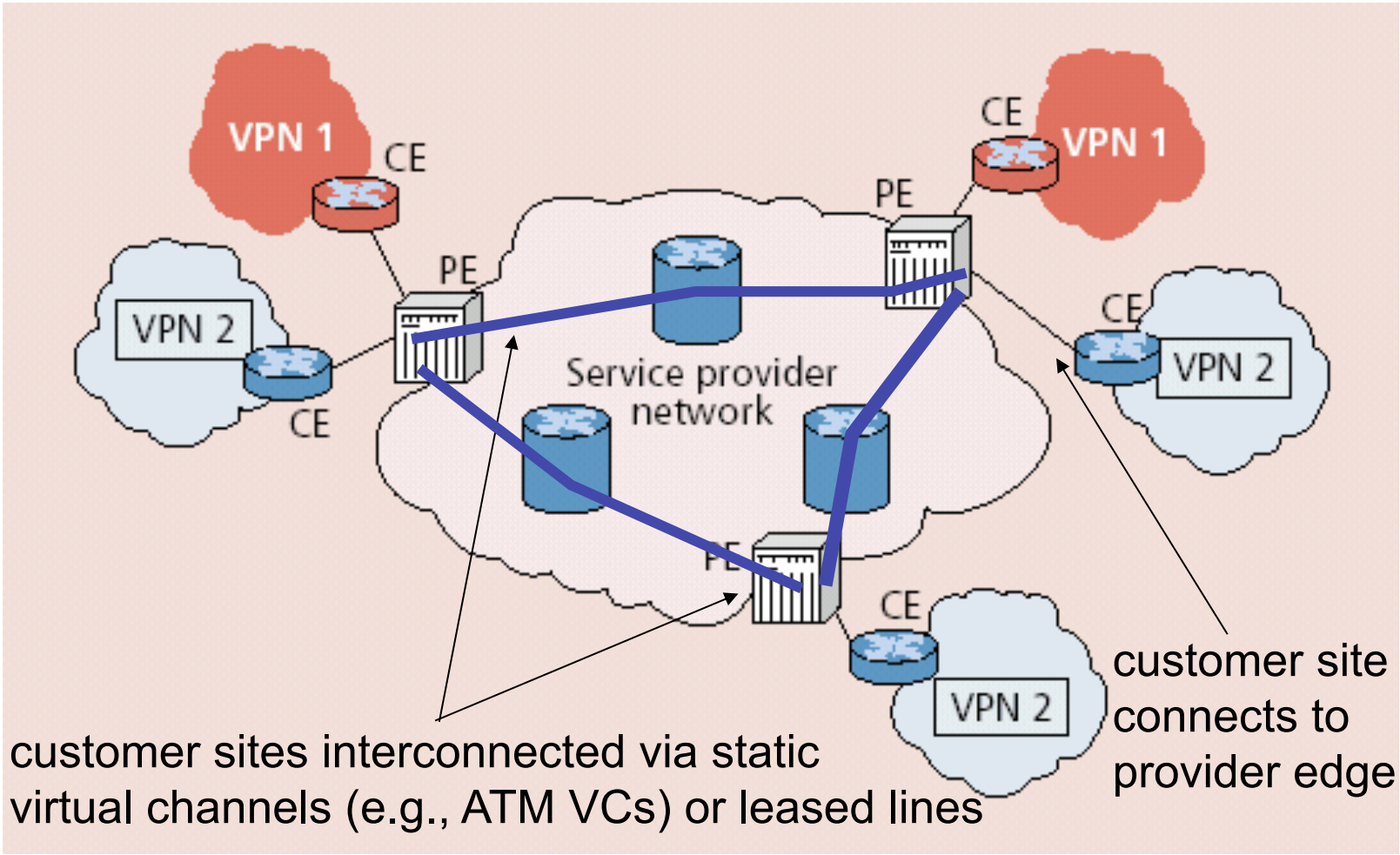
VPNs: Why?

- Privacy
- Security
- Works well with mobility (looks like you are always at home)
- Cost
  - many forms of newer VPNs are cheaper than leased line VPNs
  - ability to share at lower layers even though logically separate means lower cost
  - exploit multiple paths, redundancy, fault-recovery in lower layers
  - need isolation mechanisms to ensure resources shared appropriately
- Abstraction and manageability
  - all machines with addresses that are “in” are trusted no matter where they are

VPN: Customer Logical View

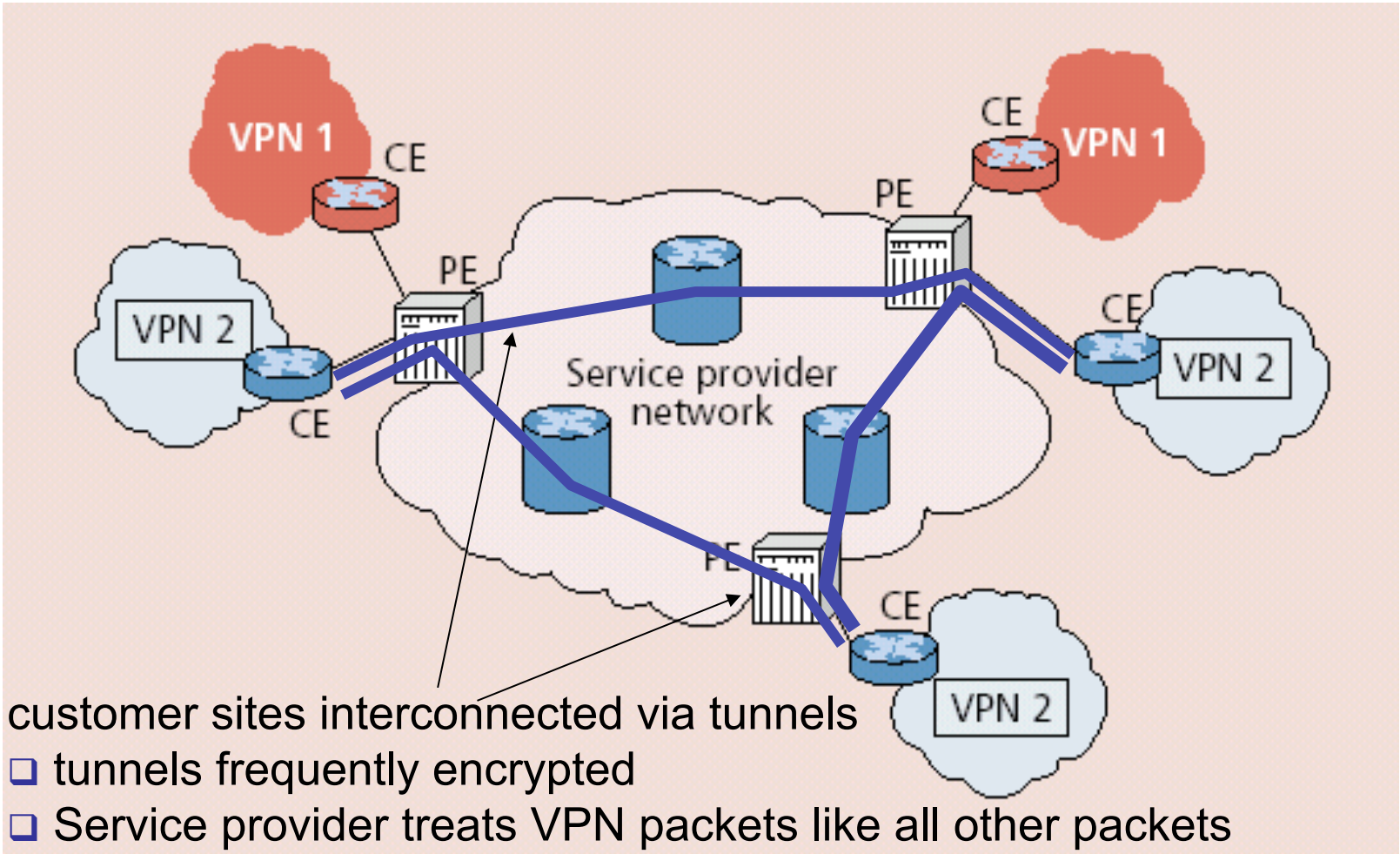


Leased-Line VPN



Customer Premise VPN

- all VPN functions implemented by customer



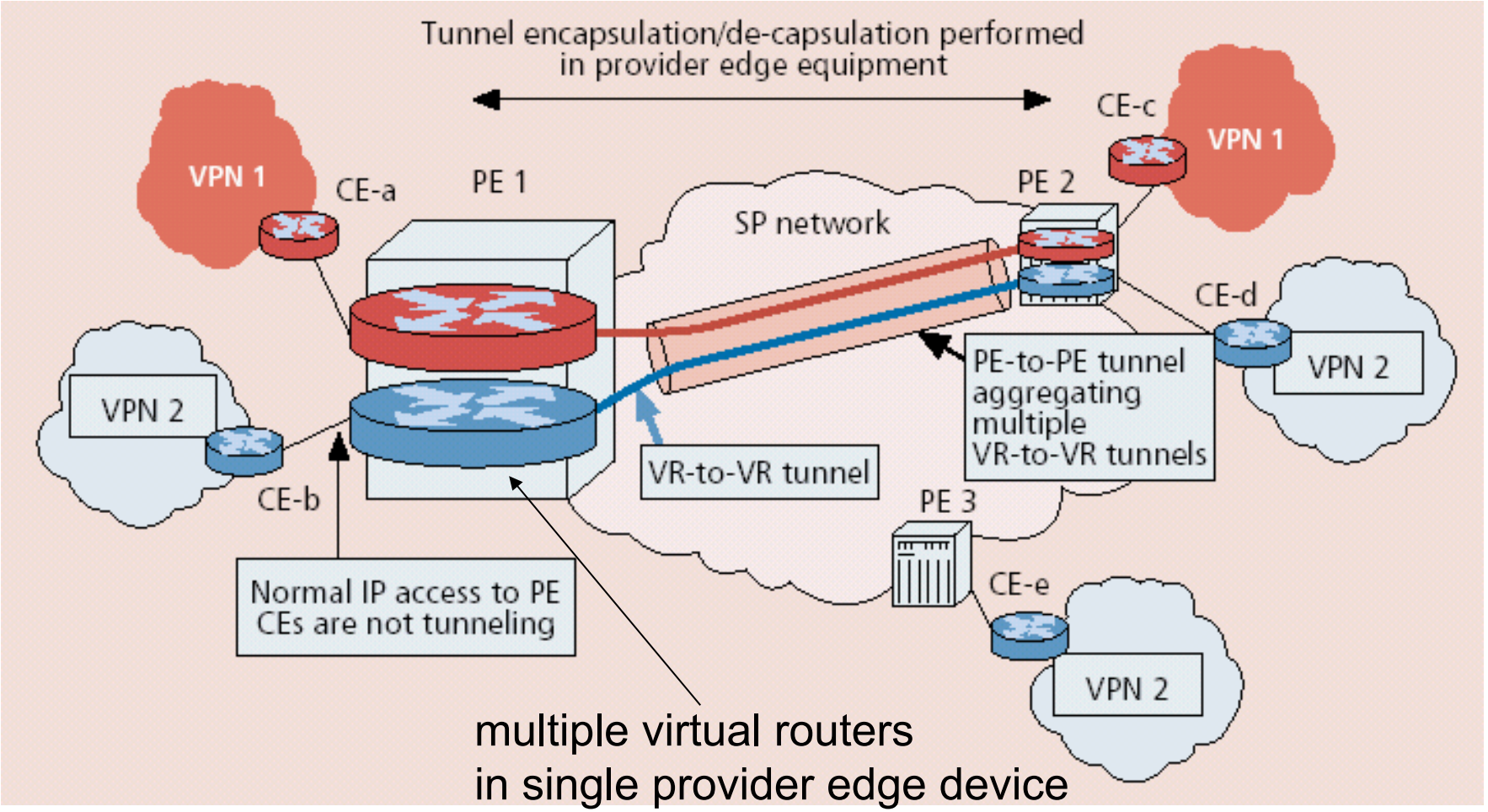
customer sites interconnected via tunnels

- tunnels frequently encrypted
- Service provider treats VPN packets like all other packets

Variants of VPNs

- Leased-line VPN
  - configuration costs and maintenance by service provider: long time to set up, manpower
- CPE-based VPN
  - expertise by customer to acquire, configure, manage VPN
- Network-based VPN
  - Customer routers connect to service provider routers
  - Service provider routers maintain separate (independent) IP contexts for each VPN
    - sites can use private addressing
    - traffic from one VPN cannot be injected into another VPN

Network-based Layer 3 VPNs



CE routers send their routes to PE routers using BGP. Routes from different VPNs remain separate in PE routers. PE routers receive IP datagrams from CE routers. Each route within a VPN is assigned a MPLS label, which is distributed by BGP.

c.f. RFC 4364: BGP/MPLS IP Virtual Private Networks (VPNs)