Computer Networks

Wenzhong Li

Nanjing University

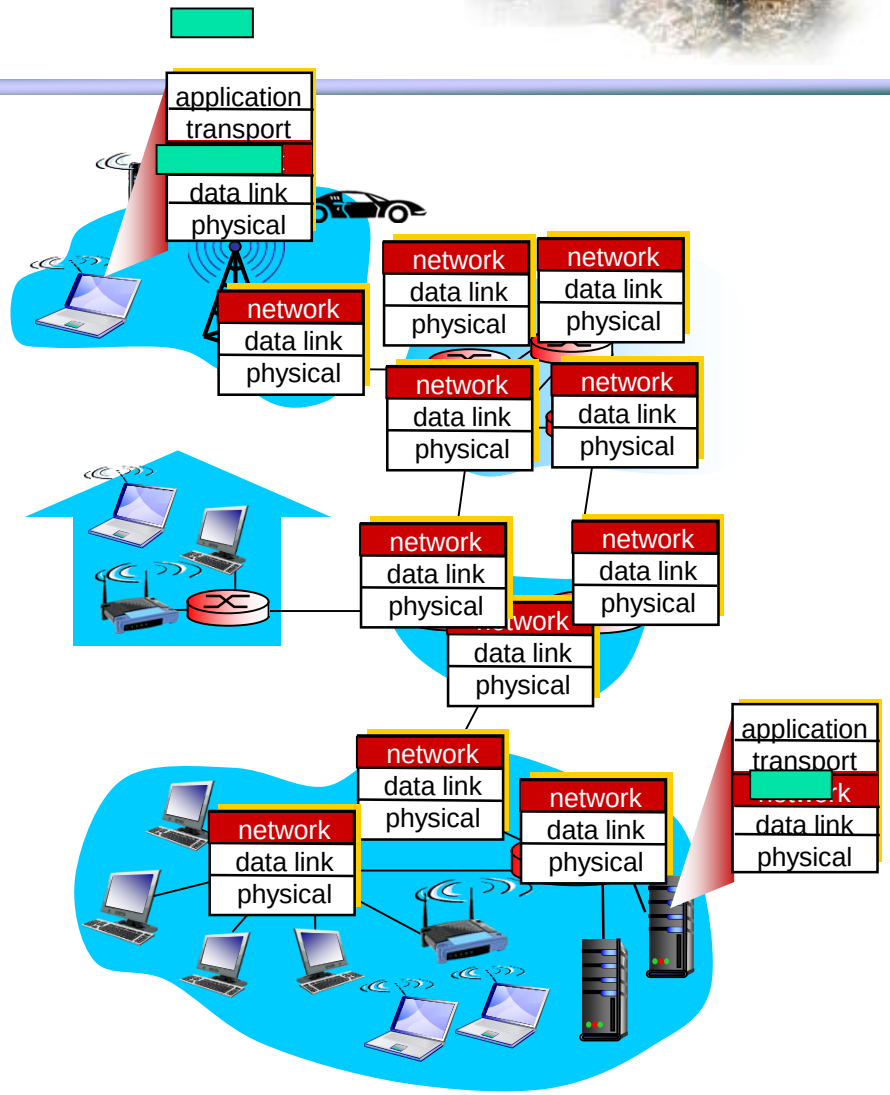
Chapter 3. Packet Switching Networks

- Network Layer Functions
- Virtual Circuit and Datagram Networks
- ATM and Cell Switching
- X.25 and Frame Relay
- Routing

Network Layer Functions

Network Layer

- transport segment from sending to receiving host
- on sending side encapsulates segments into datagrams
- on receiving side, delivers segments to transport layer
- network layer protocols in *every* host, router
- router examines header fields in all IP datagrams passing through it



Two Key Network-layer Functions

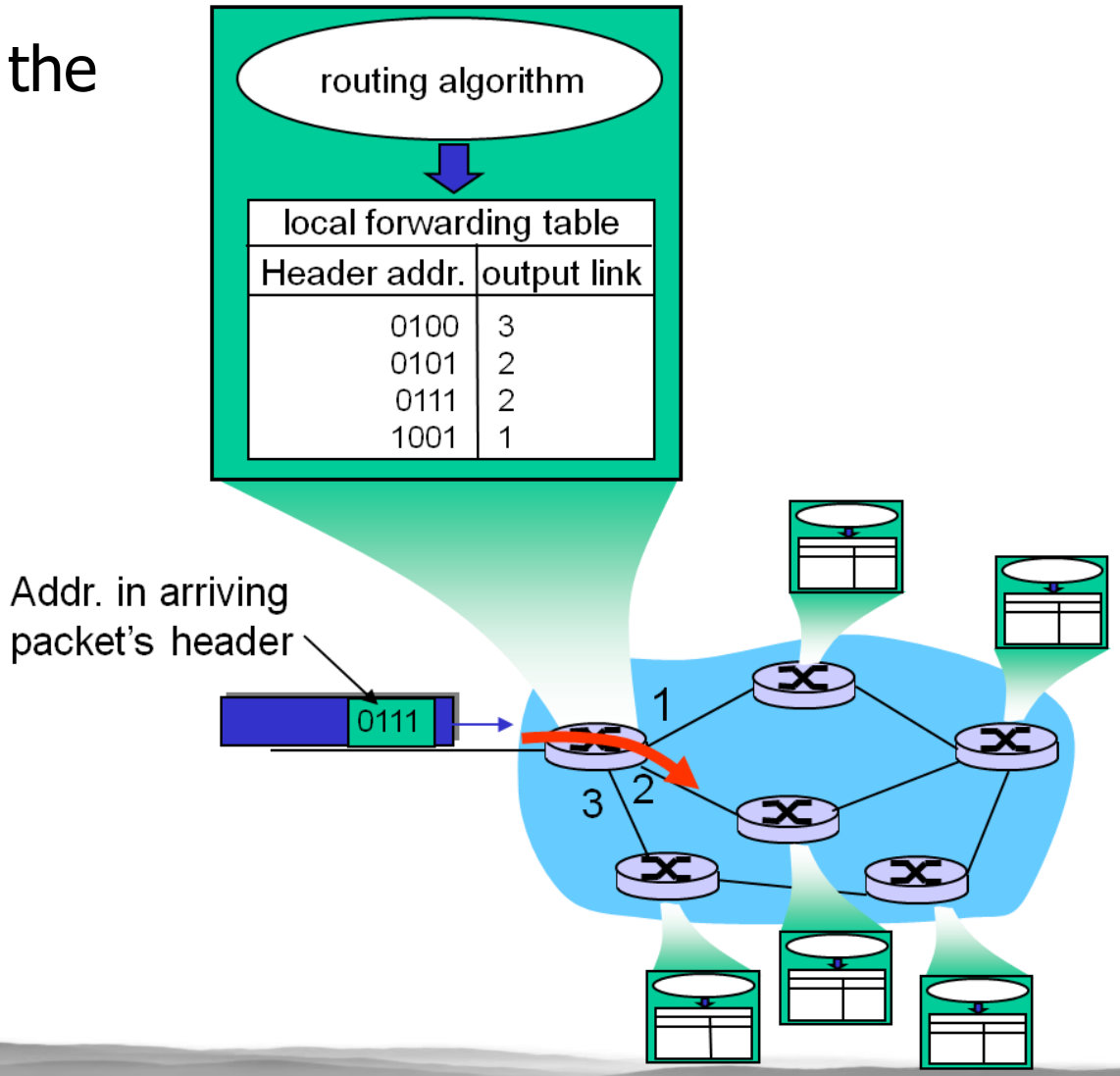
- OSI network-layer functions:
- Switching / Routing
  - Determine route taken by packets from source to destination (multiple nodes)
  - Shortest path from source to destination
  - Routing algorithms
- Forwarding
  - Move packets from input to designated output determined by switching (single node)
  - Error handling, queuing and scheduling

analogy: Trip Planning

- *routing:* planning the route from Nanjing to Shanghai (e.g., Nanjing-Wuxi-Suzhou-Shanghai)
- *forwarding:* getting through single city (e.g., entering and leaving Suzhou Station)

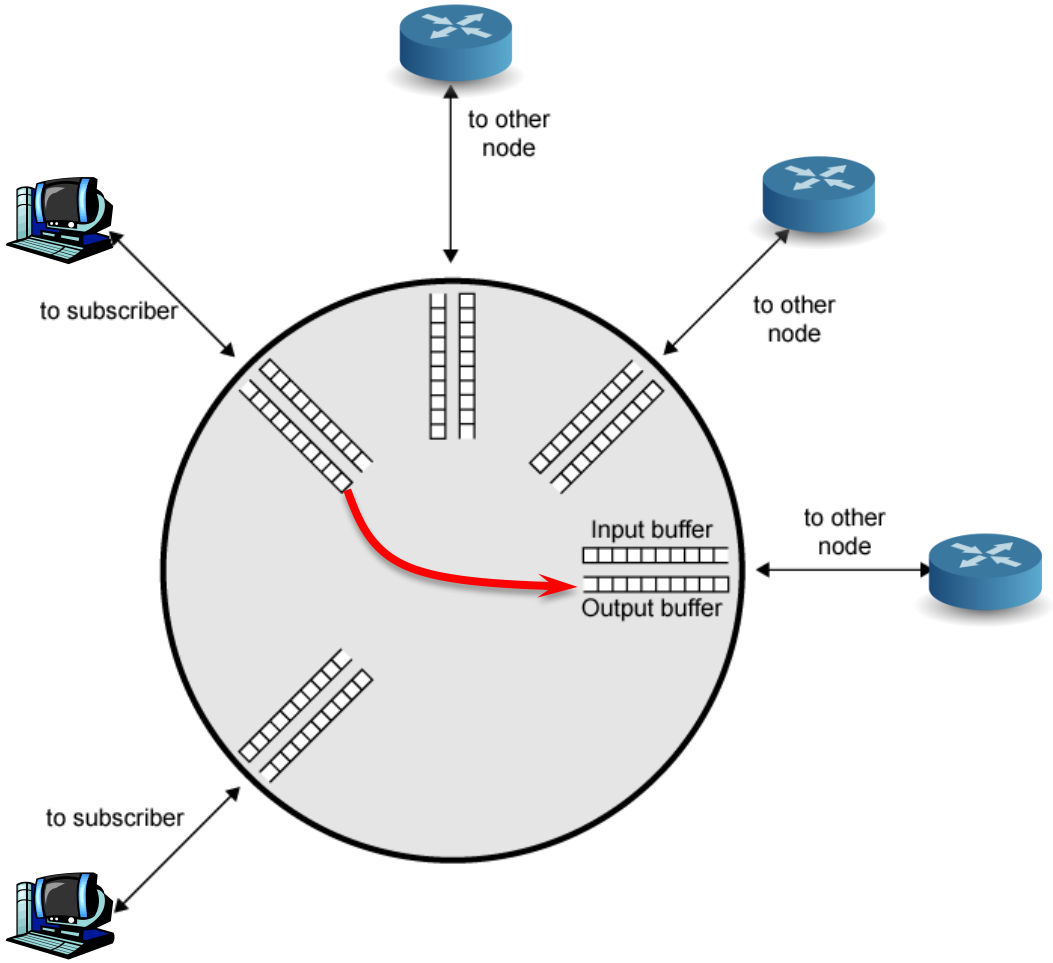
Switch Functions

- Routing determines the forwarding table



Forwarding Functions

- Queuing and scheduling
  - Host to Switch
  - Switch to Host
  - Switch to Switch



Connection setup

- 3rd important function in *some* network architectures:
  - ATM, frame relay, X.25
- Before datagrams flow, two end hosts *and* intervening routers establish virtual connection
  - Routers get involved
- Network vs transport layer connection service:
  - *network:* between two hosts (may also involve intervening routers in case of VCs)
  - *transport:* between two processes

Network Service Model

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

- Network service model
  - Service model for "channel" transporting packets from sender to receiver
  - Called Quality of Service from host perspective

Example services for individual packets

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

Example services for a flow of packets

- In-order packet delivery
- Guaranteed minimum bandwidth to flow
- Restrictions on changes in inter-packet spacing

Example：Network Service Model of ATM

In decreasing priority

- Constant Bit Rate (CBR) and Variable Bit Rate (VBR)
- Available Bit Rate (ABR) and Unspecified Bit Rate (UBR)

Network Architecture	Service Model	Guarantees ?				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

Example：Network Service Model of IP

- Best effort

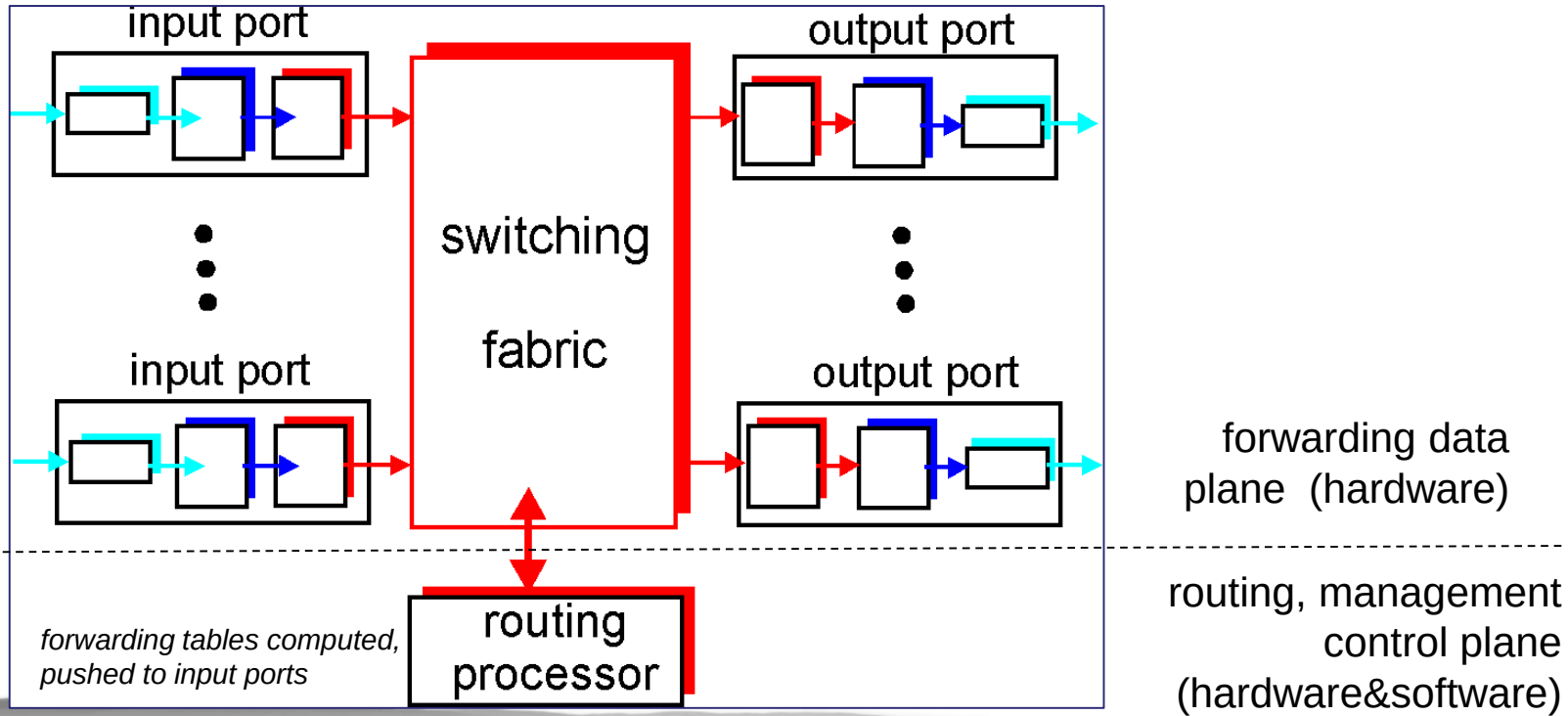
Network Architecture	Service Model	Bandwidth Guarantee	No-Loss Guarantee	Ordering	Timing	Congestion Indication
Internet	Best Effort	None	None	Any order possible	Not maintained	None
ATM	CBR	Guaranteed constant rate	Yes	In order	Maintained	Congestion will not occur
ATM	ABR	Guaranteed minimum	None	In order	Not maintained	Congestion indication provided

What's Inside a Router/Switch?

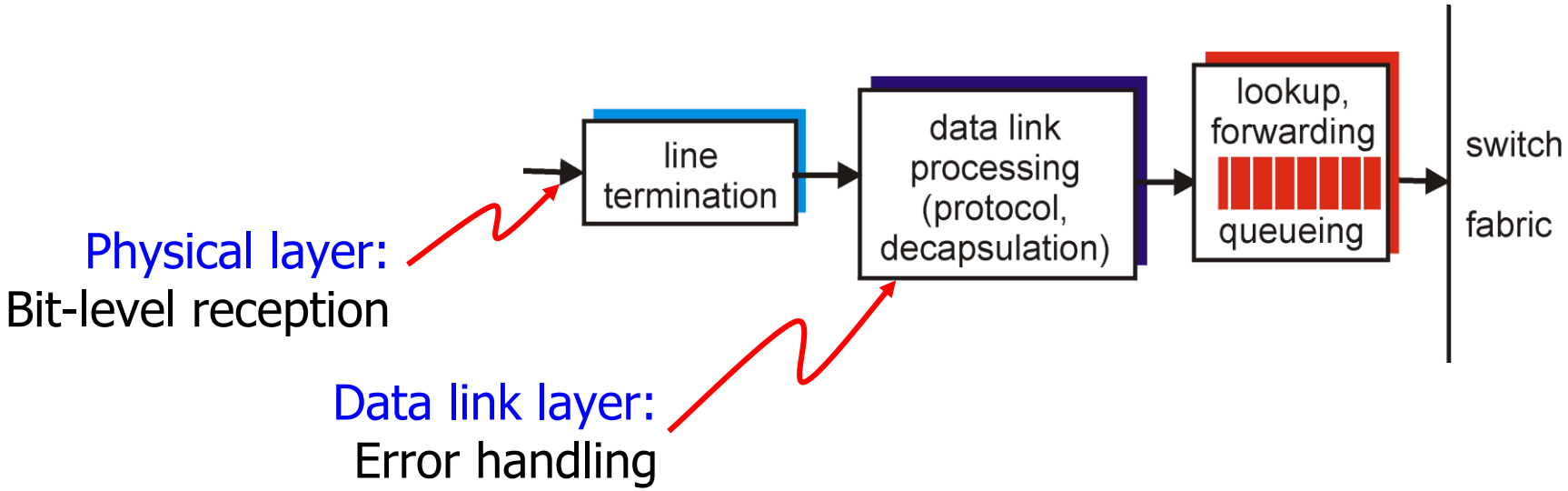
Inside a Switch: Architecture Overview

Two key switch functions:

- Run routing algorithms/protocol
- Forwarding packets from incoming to outgoing link



Input Port Functions



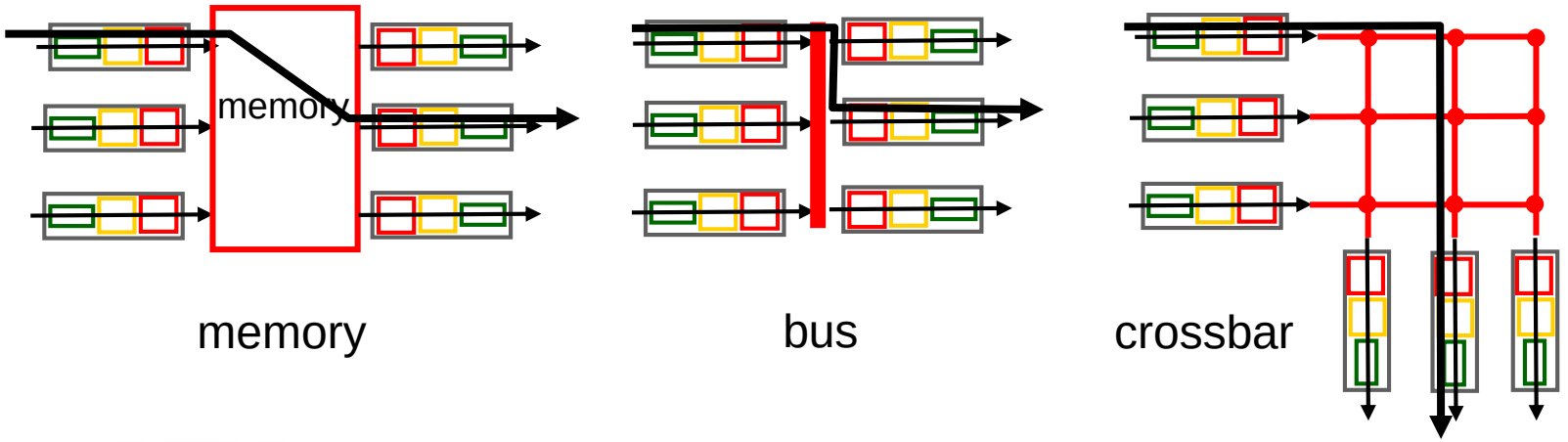
Decentralized switching

- Lookup output port using forwarding table
- Complete input port processing at "line speed"
- Queuing: if packets arrive faster than forwarding rate into switch fabric

Three Types of Switching Fabrics

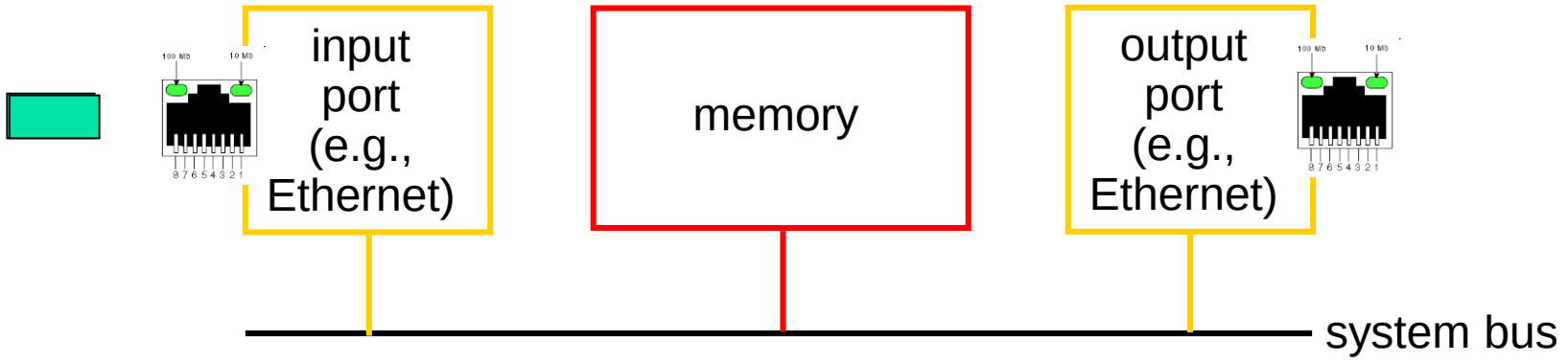
- Transfer packet from input buffer to appropriate output buffer
- Switching rate: rate at which packets can be transfer from inputs to outputs
  - often measured as multiple of input/output line rate
  - N inputs: switching rate N times line rate desirable
- Three types of switching fabrics

交换结构



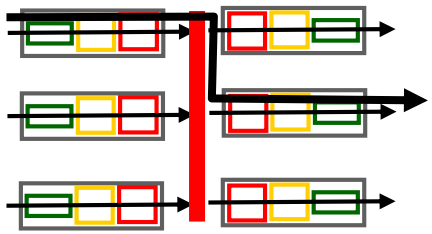
Switching via Memory

- First generation routers:
- Traditional computers with switching under direct control of CPU
- Packet copied to system's memory
- Speed limited by memory bandwidth (2 bus crossings per datagram)



Switching via a Bus

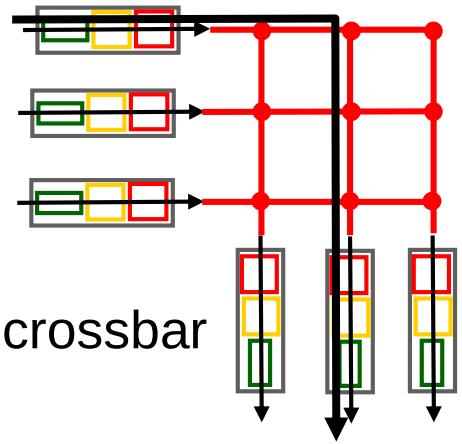
- Datagram from input port memory to output port memory via a shared bus
- *Bus contention:* switching speed limited by bus bandwidth
- 32 Gbps bus, Cisco 5600: sufficient speed for access and enterprise routers



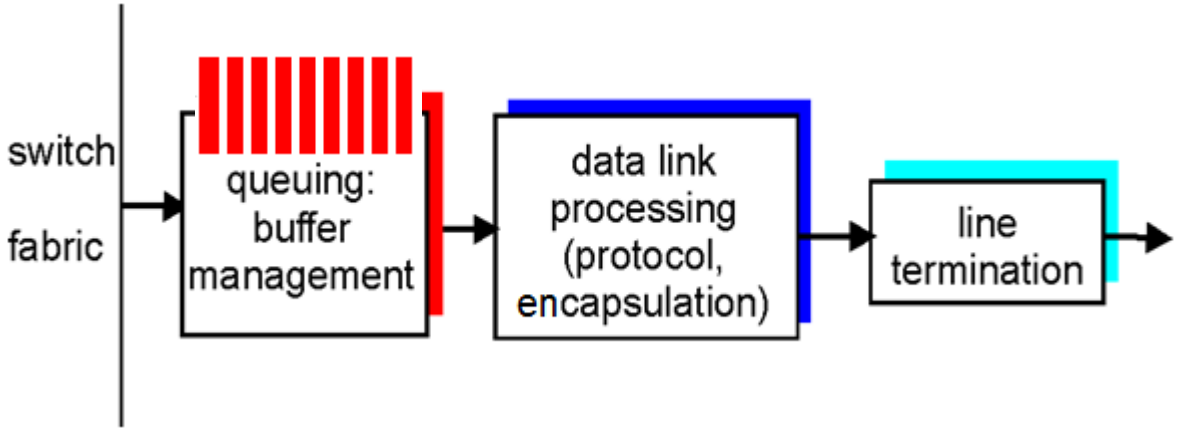
bus

Switching via a Mesh

- Overcome bus bandwidth limitations
- Banyan networks, crossbar, other interconnection nets initially developed to connect processors in multiprocessor
- Advanced design: fragmenting datagram into fixed length cells, switch cells through the fabric.
- Cisco 12000: switches 60 Gbps through the interconnection network



Output Port Functions



- Buffering
  - Required when packets arrive from fabric faster than the transmission rate

Datagram (packets) can be lost due to congestion, lack of buffers

- Scheduling discipline
  - Chooses among queued packets for transmission
  - Select packets to drop when buffer saturates

Priority scheduling – who gets best performance

How much buffering?

- RFC 3439 rule of thumb: average buffering equal to "typical" RTT (say 250 msec) times link capacity C
  - e.g., C = 10 Gpbs link: 2.5 Gbit buffer

- Recent recommendation: with *N* flows, buffering equal to [Appenzeller 2004]

$$\frac{RTT \cdot C}{\sqrt{N}}$$

Virtual Circuit and Datagram Networks

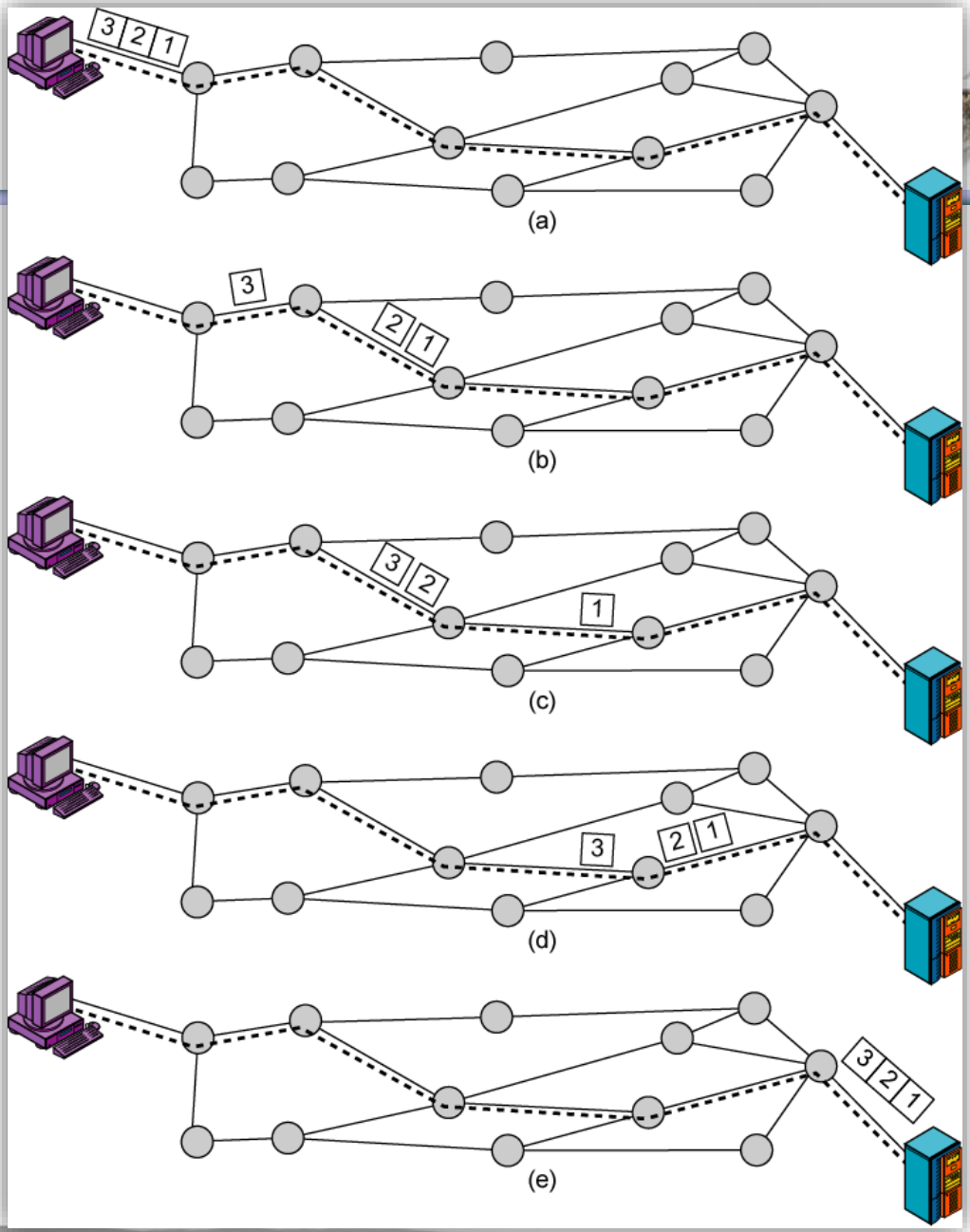
Recap: Circuit Switching & Packet Switching

- Circuit Switching
  - End-to-end resources reserved for "call"
    - Link bandwidth, switch capacity
  - Dedicated resources: no sharing
  - Guaranteed performance
  - Call setup/teardown required
- Packet Switching
  - Each end-to-end data stream divided into packets
  - Application A, B packets share network resources
  - Store and forward: packets move one hop at a time, stored (queued) at switches
  - Resource contention: aggregate (burst-up) resource demand can exceed amount available
  - Congestion: packets queue and wait for link use

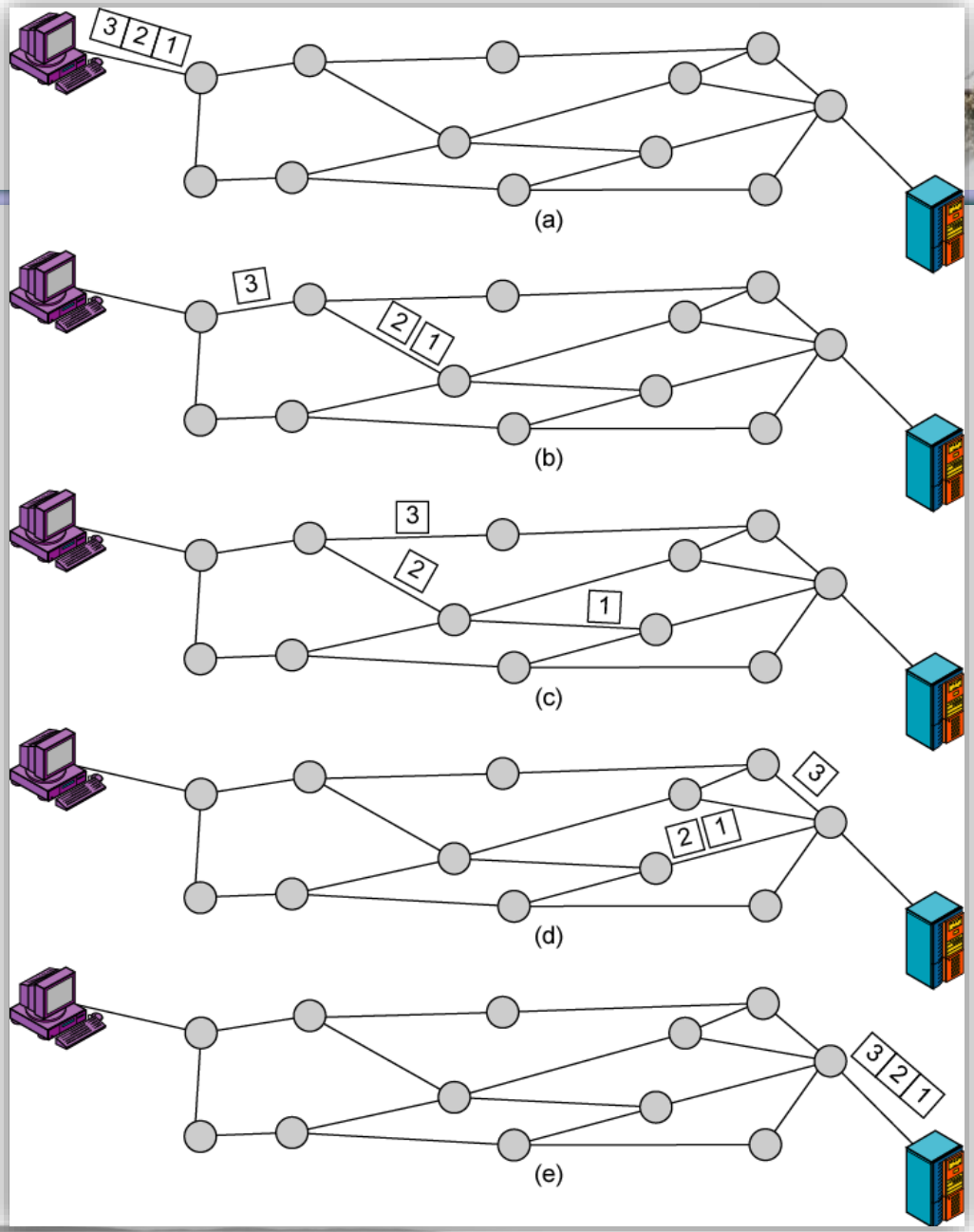
Virtual Circuit and Datagram Networks

- Two types of Package Switch Networks
  - Virtual circuit networks
    - Network service provided on flow of packets
    - VC network provides network-layer connection oriented service
    - E.g., ATM, X.25, Frame Relay
  - Datagram networks
    - Network service provided on singular packet
    - Datagram network provides network-layer connectionless service
    - E.g., IP network

Routing in Virtual Circuit



Routing in Datagram Nets



Virtual Circuit Networks

- Connection setup, teardown for each flow of packets
- Each packet carries VC identifier (not destination host address)

- Every switch on source-destination path maintains "state" for each passing connection
- Link, switch resources (bandwidth, buffers) may be allocated to VC
  - Dedicated resources = predictable quality of service

Connection Setup

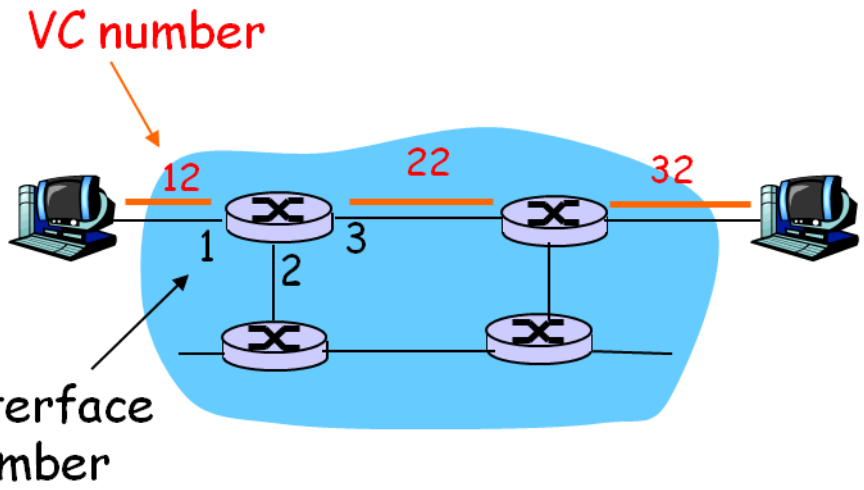
- Essential function for virtual circuit networks
  - E.g. ATM, frame relay, X.25
- Two end hosts and intervening switches pre-establish a path for virtual connection
- Routing is used for finding a suitable (shortest) path

VC Implementation

- A VC consists of
  - Path from source to destination
  - VC numbers, maybe one number for each link along the path
  - Entries in forwarding tables in switches along the path

- Note:
  - Packet belonging to VC carries VC number (rather than addresses)
  - VC number can be changed on each link, forwarding table lists the new VC number

A Forwarding Table for VC



Forwarding table in northwest switch

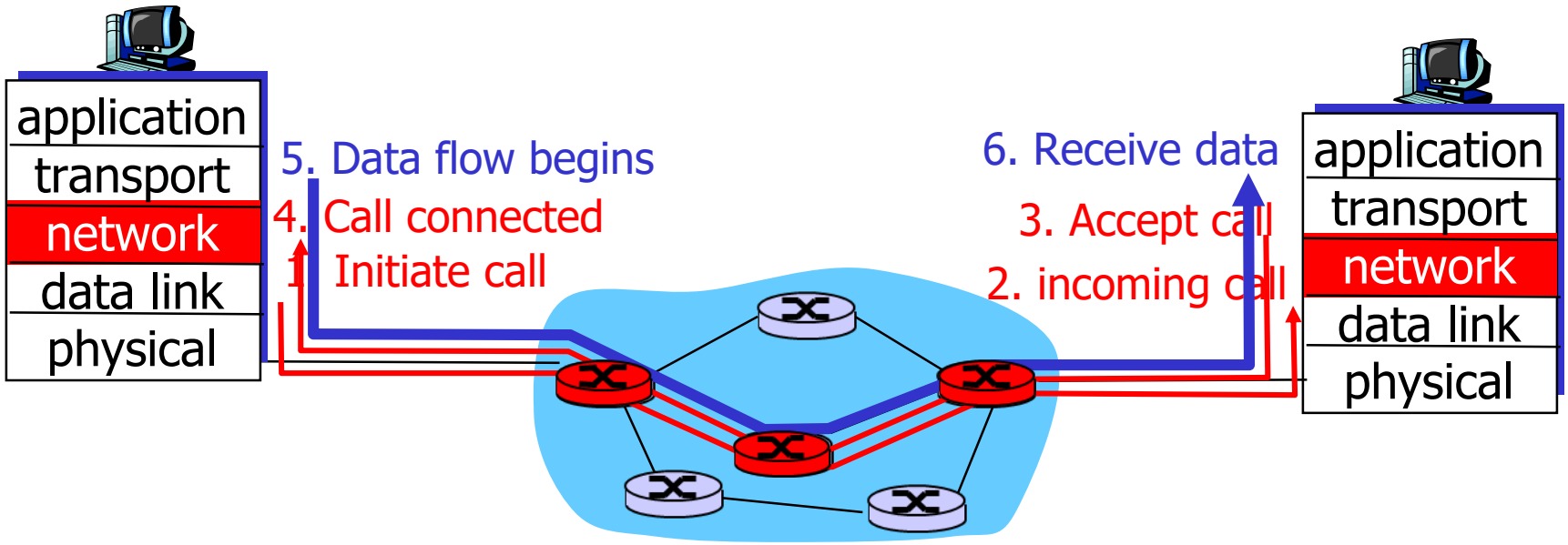
Incoming interface	Incoming VC #	Outgoing interface	Outgoing VC #
1	12	3	22
2	63	1	18
3	7	2	17
1	97	3	87
…	…	…	…

Table entries constitutes state information of a VC

Virtual Circuits: Signaling Protocols

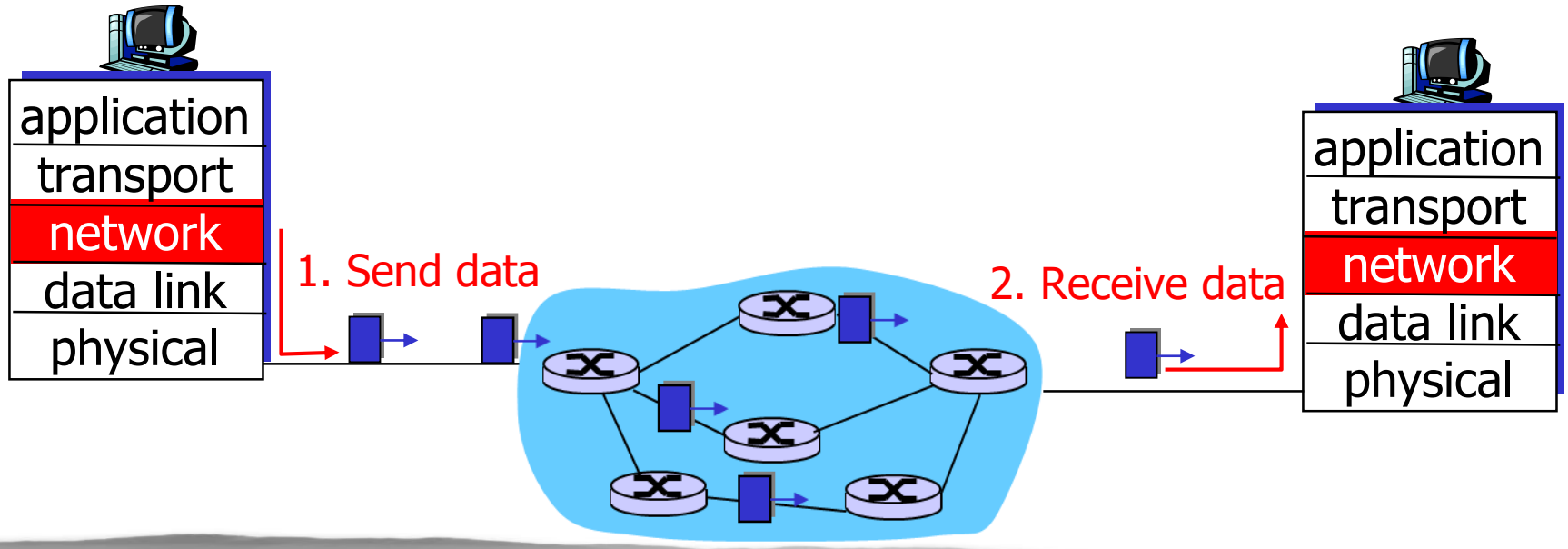
- Used to setup, maintain and teardown VC
- Used in ATM, frame-relay, X.25
- Not used in today's Internet

信令协议



Datagram Networks

- No call setup at network layer
- No network-level concept of "connection"
- Switches: no state about end-to-end connections
- Packets forwarded using destination host address
- Packets between same source-dest pair may take different paths



A Forwarding Table for Datagram Networks

- Also called routing table
- May reach 4 billion entries

- The destination address prefix may define a switch address or a subnet address

Dest Address Prefix	Address Mask	Link Interface
11001000 00010111 00010	11111111 11111111 11111000 00000000	0
11001000 00010111 00011000	11111111 11111111 11111111 00000000	1
11001000 00010111 000110	11111111 11111111 11111100 00000000	2
default	*	3

Longest Prefix Matching

Address Prefix	Link Interface
11001000 00010111 00010	0
11001000 00010111 00011000	1
11001000 00010111 00011	2
otherwise	3

Examples

DA: 11001000 00010111 00010110 10100001 Which interface?

DA: 11001000 00010111 00011000 10101010 Which interface?

Longest prefix matching rule: when looking for forwarding table entry for given destination address, use longest address prefix that matches destination address.

Datagram vs. Virtual Circuit

Datagram (Internet)

- Data exchange among computers
  - "Elastic" service, no strict timing
- "Smart" end systems (computers)
  - Can adapt, perform control, error recovery
  - Simple inside network, complexity at "edge"
- Many link types
  - Different characteristics
  - Uniform service difficult

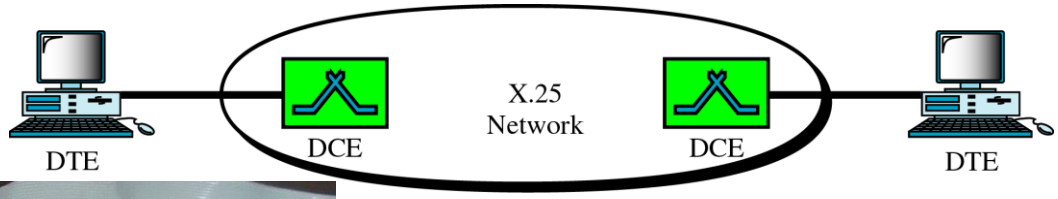
Virtual Circuit (ATM)

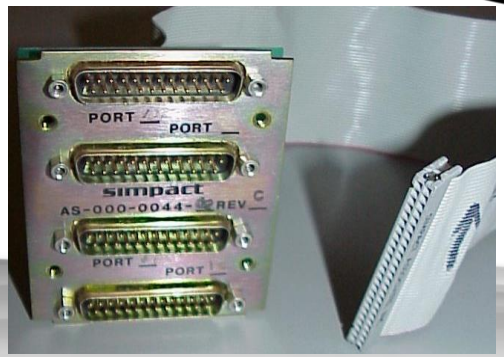
- Evolved from telephony
- Human conversation:
  - Strict timing, reliability requirements
  - Need guaranteed service
- "Dumb" end systems
  - Telephones
  - Complexity inside network (switches)
- Link type standardized

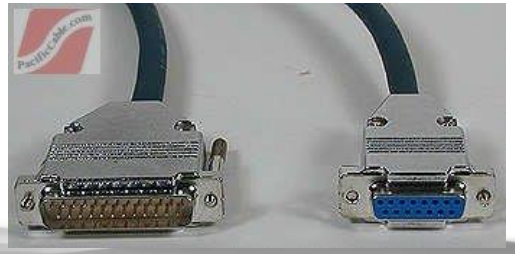
X.25, Frame Relay, and ATM

X.25

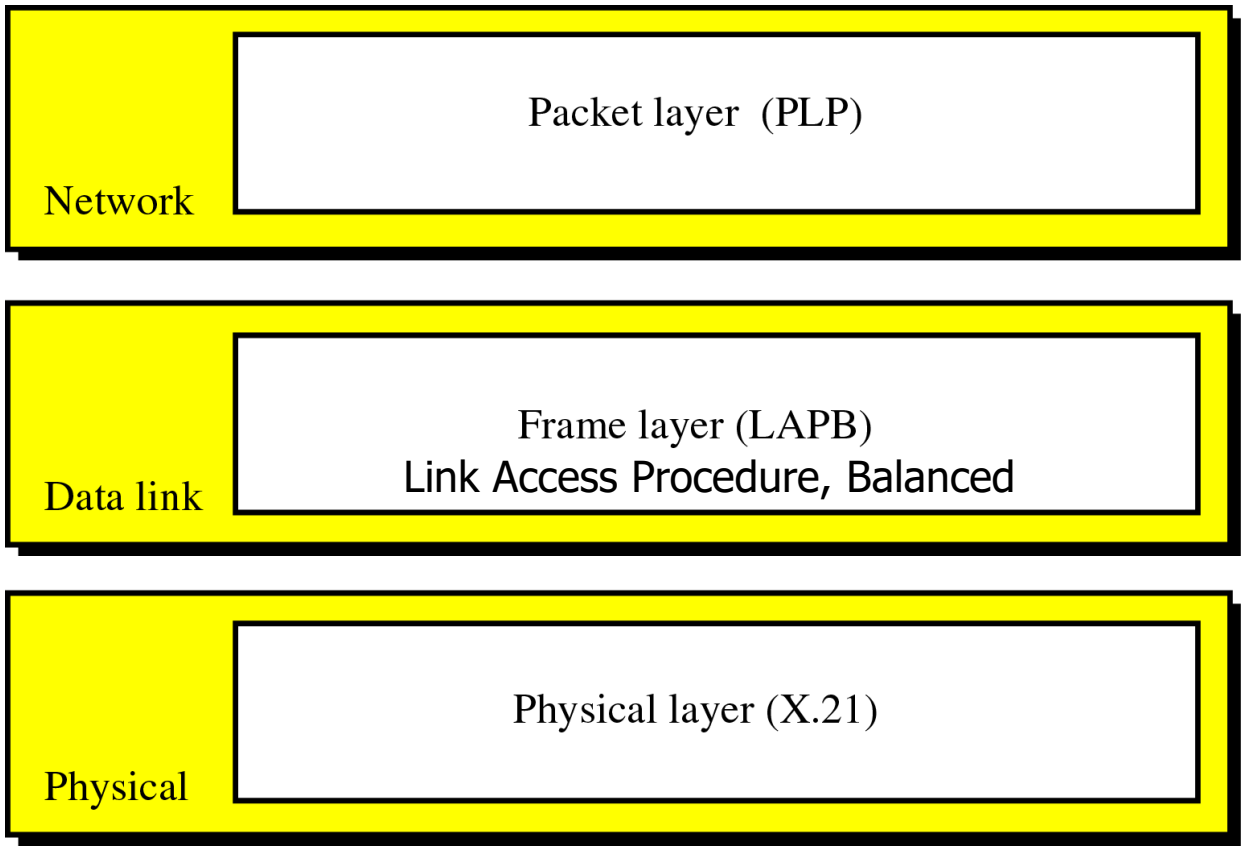
- A packet-switching wide area network developed by ITU-T in 1976 （第一个面向连接的网络，第一个公共数据网络，一种使用电话或者ISDN设备作为网络硬件设备来架构广域网的ITU-T网络协议）
- Defines how a packet-mode terminal can be connected to a packet network
- Defines how a user's DTE (Data Terminal Equipment) communicates with DCEs (Data Communications Equipments) in a packet switching network
- Defines how packets are sent thru the virtual circuit established between DTEs







X.25 Layers



PLP: Packet Layer Protocol

Frame Relay

- Frame Relay
  - Packet-switching with virtual-circuit technology
  - An enhancement of X.25, due to improved transmission media
  - Interconnect LANs, instead of terminals
- Improvement of X.25, taking advantage of high-speed new links with lower error-rates
  - Operate only at the Physical and Data link layer（提供数据链路层和物理层的协议规范，任何高层协议都独立于帧中继协议）
  - Not provide error checking or require ACK in data link layer
- Layers in FR
  - Physical layer, any protocols recognized by ANSI, up to 44.376 Mbps
  - Data link layer, a simplified version of HDLC called core LAPF, no error and flow control fields

ANSI: American National Standards Institute
LAPF: Link Access Procedure for Frame Mode Services

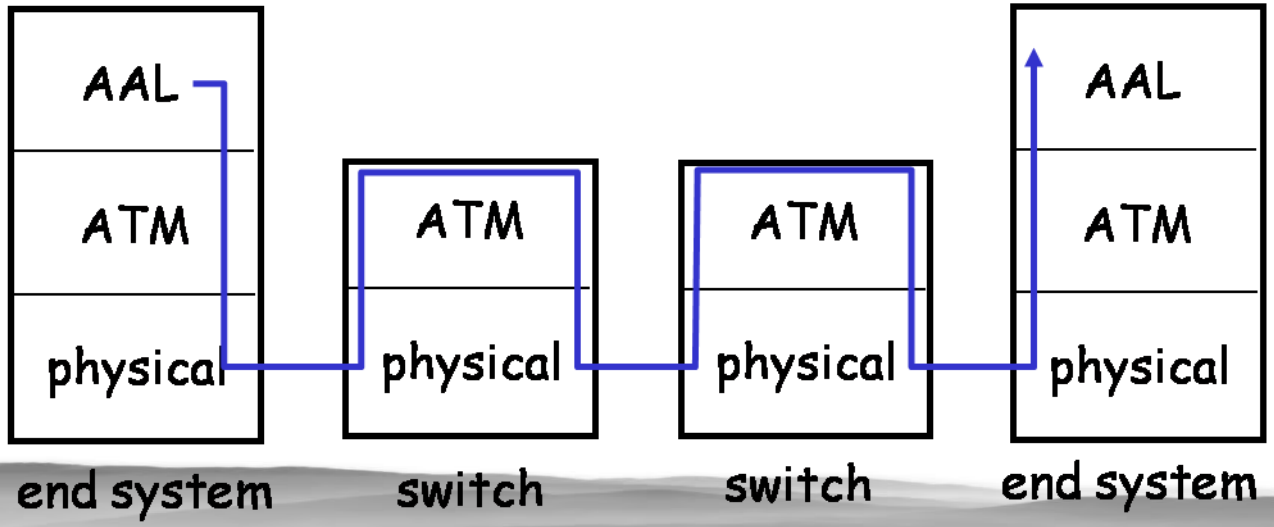
ATM and Cell Switching

- ATM: Asynchronous Transfer Mode
  - 1990's/2000 standard for high-speed Broadband Integrated Service Digital Network (ISDN, 综合业务数字网) architecture
  - 155Mbps to 622 Mbps and higher

- Features
  - Meeting timing/QoS requirements of voice and video, also support "burst" data
  - "Next generation" telephony: technical roots in telephone world
  - Packet-switching (fixed length packets, called "cells") using virtual circuits

ATM Architecture

- Adaptation layer: only at edge of ATM network
  - Data segmentation/reassembly, different service models
  - Roughly analogous to Internet transport layer
- ATM layer: "network" layer
  - Cell switching, routing
- Physical layer: SDH/SONET



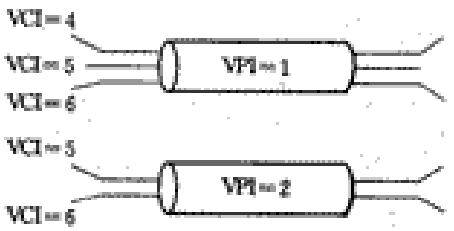
ATM Adaptation Layer

- ATM Adaptation Layer (AAL)
  - "Adapts" upper layers (IP or native ATM applications) to ATM layer below
  - Present only in end systems, not in switches
- Different types of AALs
  - AAL1, Constant Bit Rate (CBR), e.g. circuit emulation
  - AAL2, Variable Bit Rate (VBR), e.g. voice and video
  - AAL3/4, Connection-oriented data service, e.g. X.25 and Frame Relay
  - AAL5, Connectionless data service, e.g IP datagram

ATM Layer: Virtual Circuits

- VC transport: cells carried on VC from source to destination
- Permanent VCs (PVC)
  - Long lasting connections
- Switched VCs (SVC)
  - Dynamically set up on per-connection basis
- A VC consists of virtual paths and virtual channels
  - Virtual Path Identifier (VPI) + Virtual Channel Identifier (VCI)

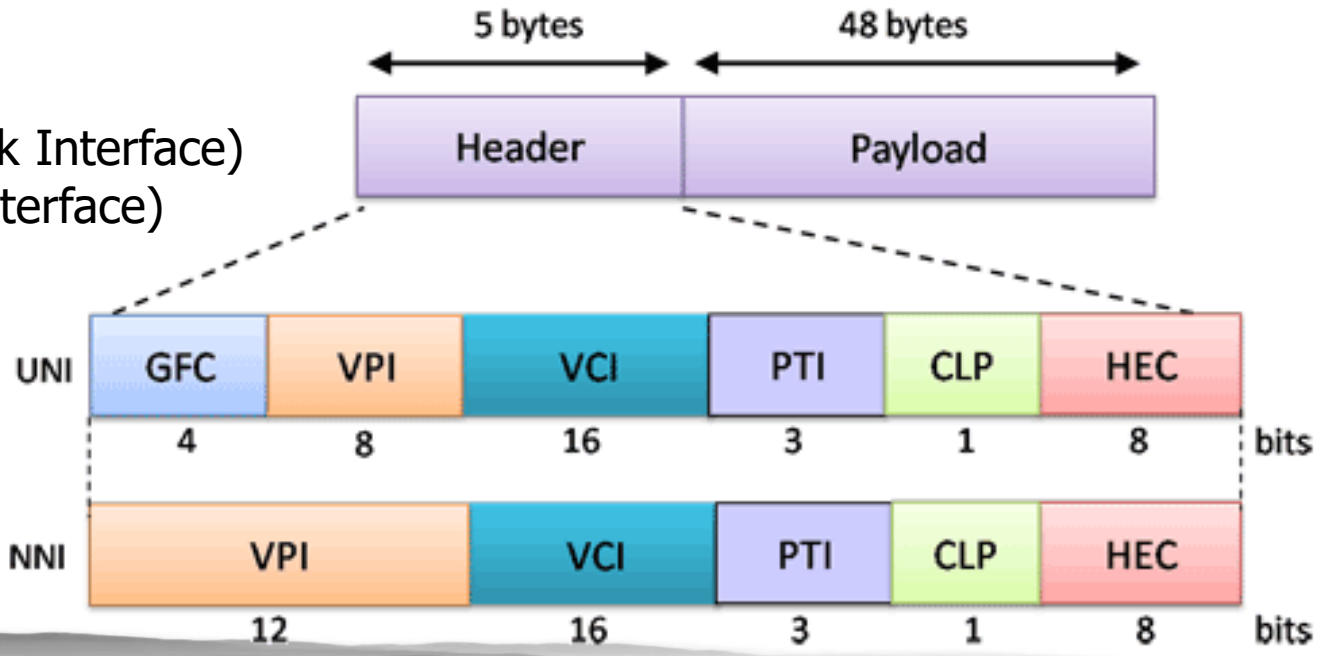
信元路由信息：
虚通道(virtual paths)：由VPI指定，一个VPI包含一组VCI
虚通路(virtual channels)：由VCI指定



ATM Cells

- 5 octet header + 48 octet payload
- Small payload → short cell-creation delay and switching delay
- 48 = halfway between 32 (Europe) and 64 (North America), a compromise

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



ATM Physical Layer

- 2 sublayers
  - Transmission Convergence (TC) sublayer
    - Header checksum generation: 8 bits CRC
    - Cell delineation to signal representation
    - Transmission of idle cells when no data cells to send
  - Physical Medium Dependent (PMD) sublayer: depends on physical medium being used
    - E.g., SONET/SDH, TI/T3, etc

传输聚合子层:
在发送方,它从ATM层接收信元,组装成特定形式的帧(SONET帧或FDDI数据帧). 在接收方,它从PMD子层提取信元,交付ATM层. 类似于链路层功能
物理介质相关子层：指定物理特性

Summary

- 网络层基本功能
  - 交换/路由，转发，建立连接
- 路由器的构成
- 两种分组交换网络
  - 虚电路网络
    - ATM（面向连接，信元：固定长度的分组，支持CBR， VBR， ABR， UBR）
    - X.25（面向连接，流控制和错误检测），帧中继（面向连接，无错误控制，无流控制）
  - 数据报网络
    - IP网络

Homework

- 第四章：R1, R2, P1, P2