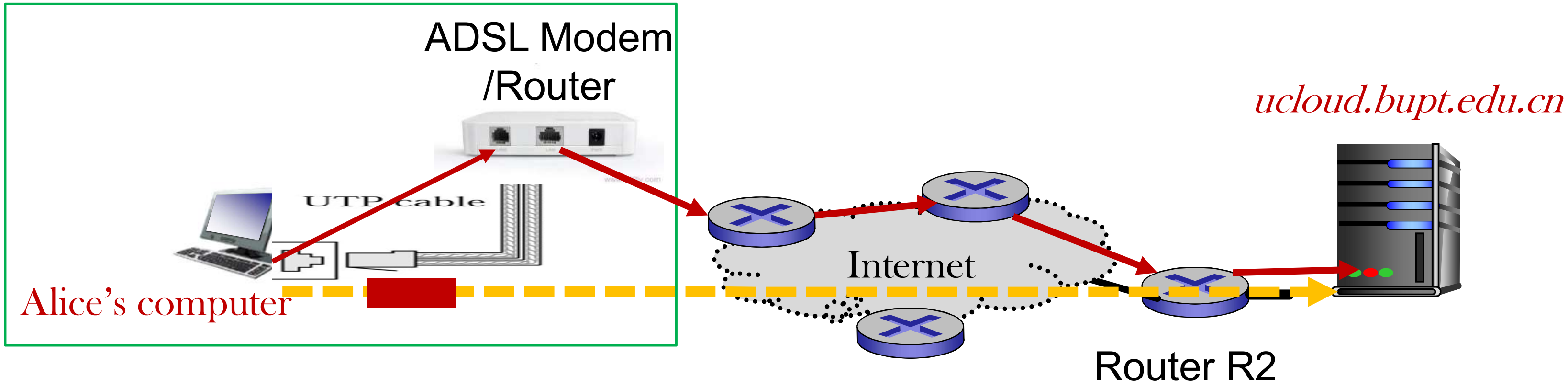
Chapter5
The Network Layer

(around 14 hours)

The task: Uploading homework image to 教学云平台



- WWW requirements
  - Environment: internetworking
- Function of Data Link Layer
  - Environment: between neighbors (host-router, router-router)
- How to find a path from source to destination?

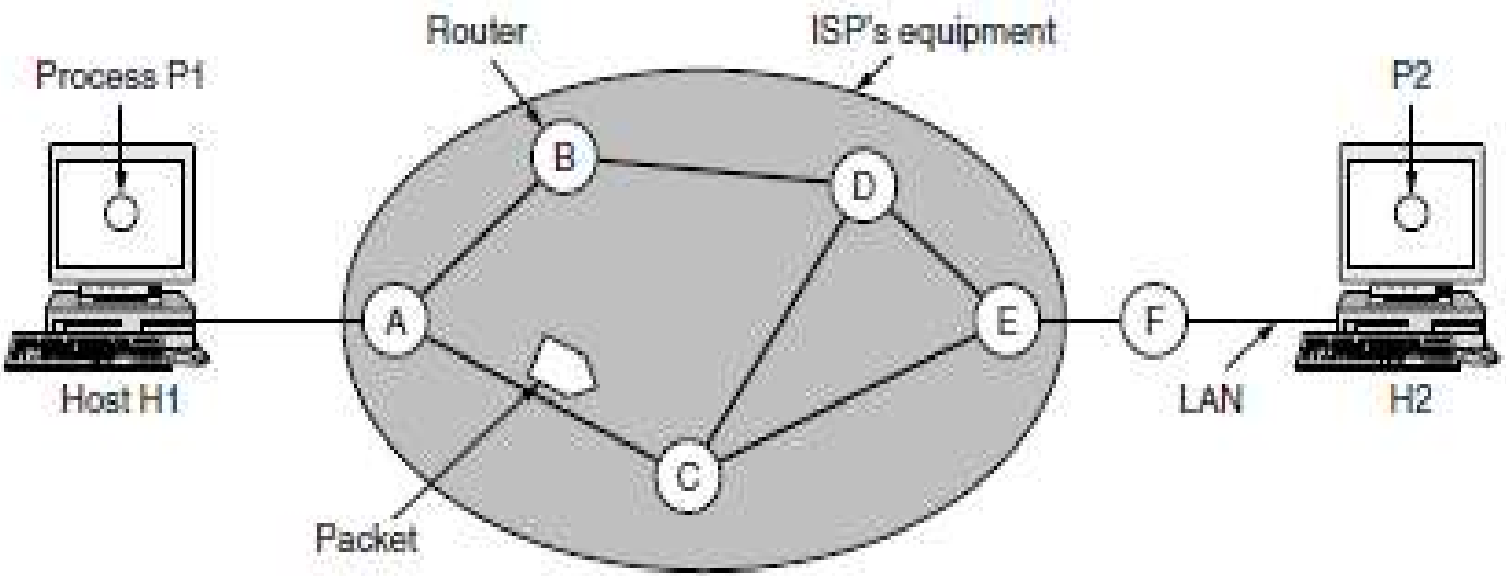
→Network layer

Outline

- *Introduction to Network Layer*
  - Position and functions of network layer
  - Virtual Circuit Switching and Datagram Routing
- Routing Algorithms
  - Distance Vector Routing (DVR)
  - Link State Routing (LSR)
  - Hierarchical Routing
- Congestion Control and Internetworking
- The network layer in Internet
  - IP Addressing
  - Packet format
  - IPv6
  - ICMP、ARP and DHCP
  - Routing Protocols: RIP, OSPF and BGP
  - IP Multicast & Mobile IP

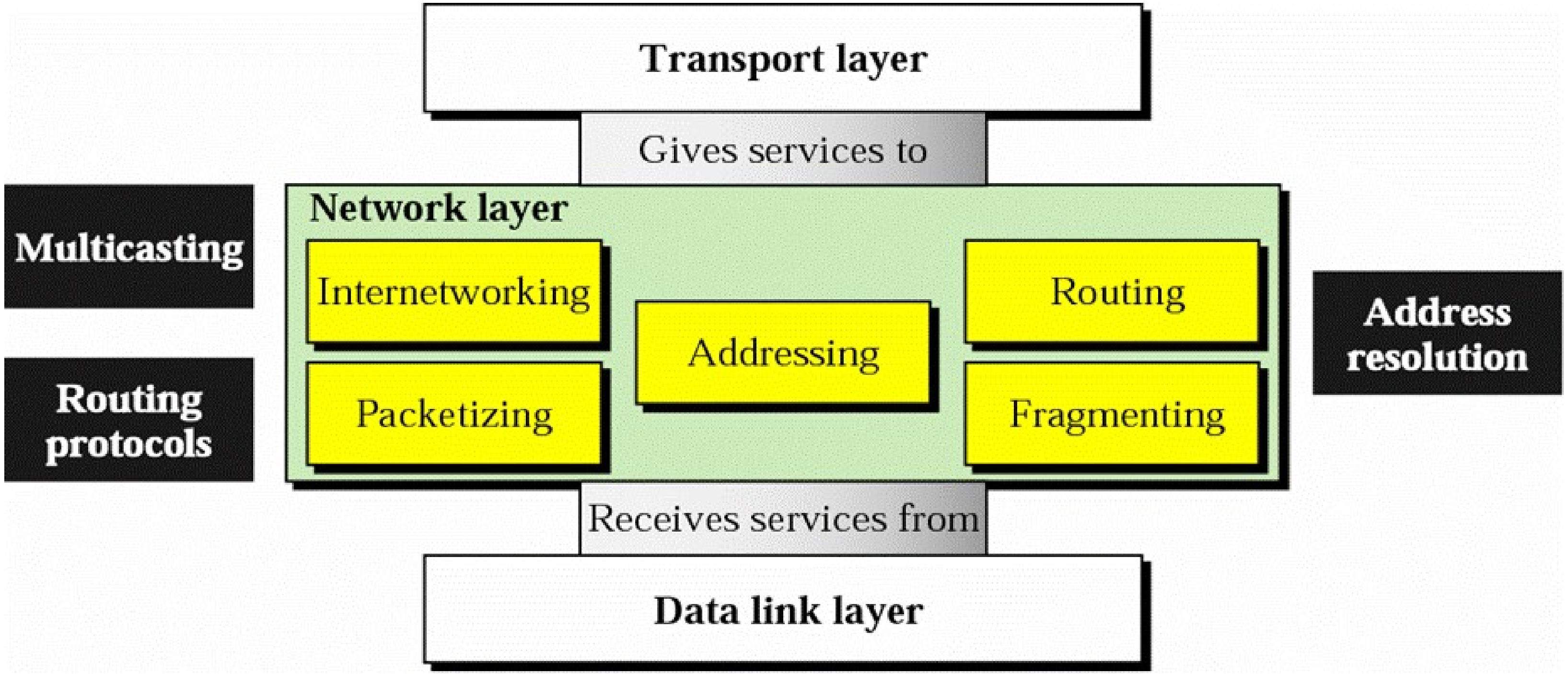
Environment of Network Layer Protocols

网络互连：Interconnection of sub networks(子网) by routers



Store-and-Forward Packet Switching
(存储转发分组交换)

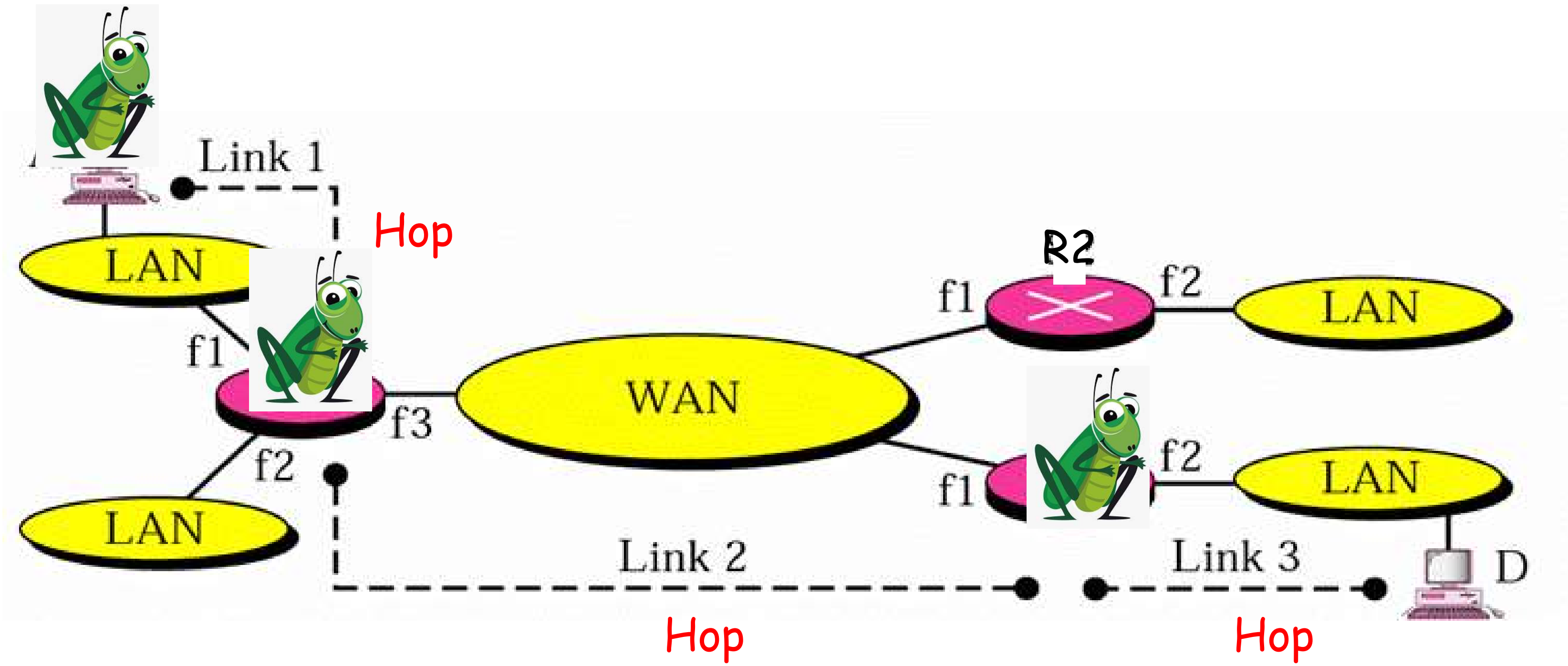
Positions and Functions of Network Layer



Network layer is the lowest layer that deals with end-to-end (host-to-host) transmission (端到端/主机-主机传输)

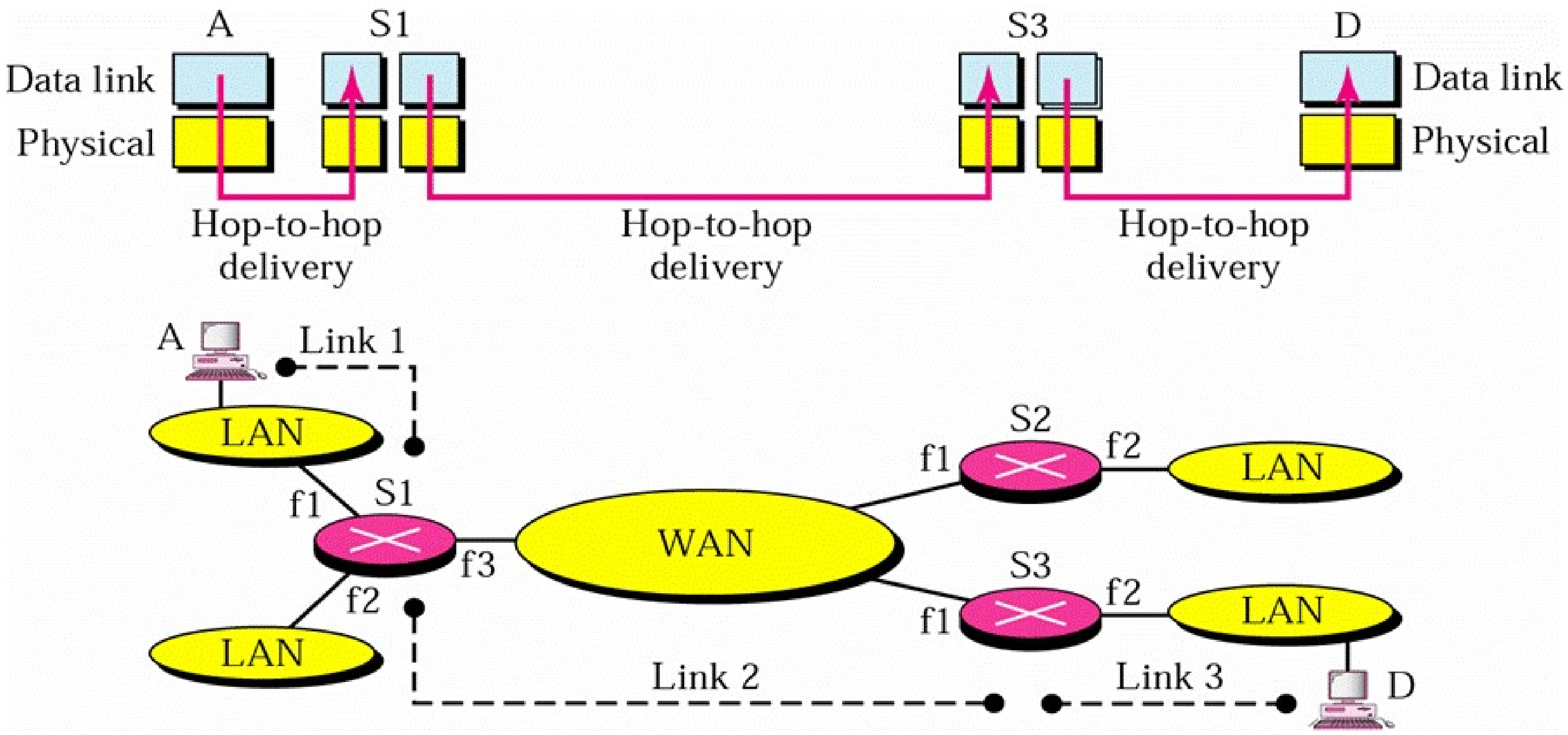
[Forouzan]

The concept of Hop(跳)



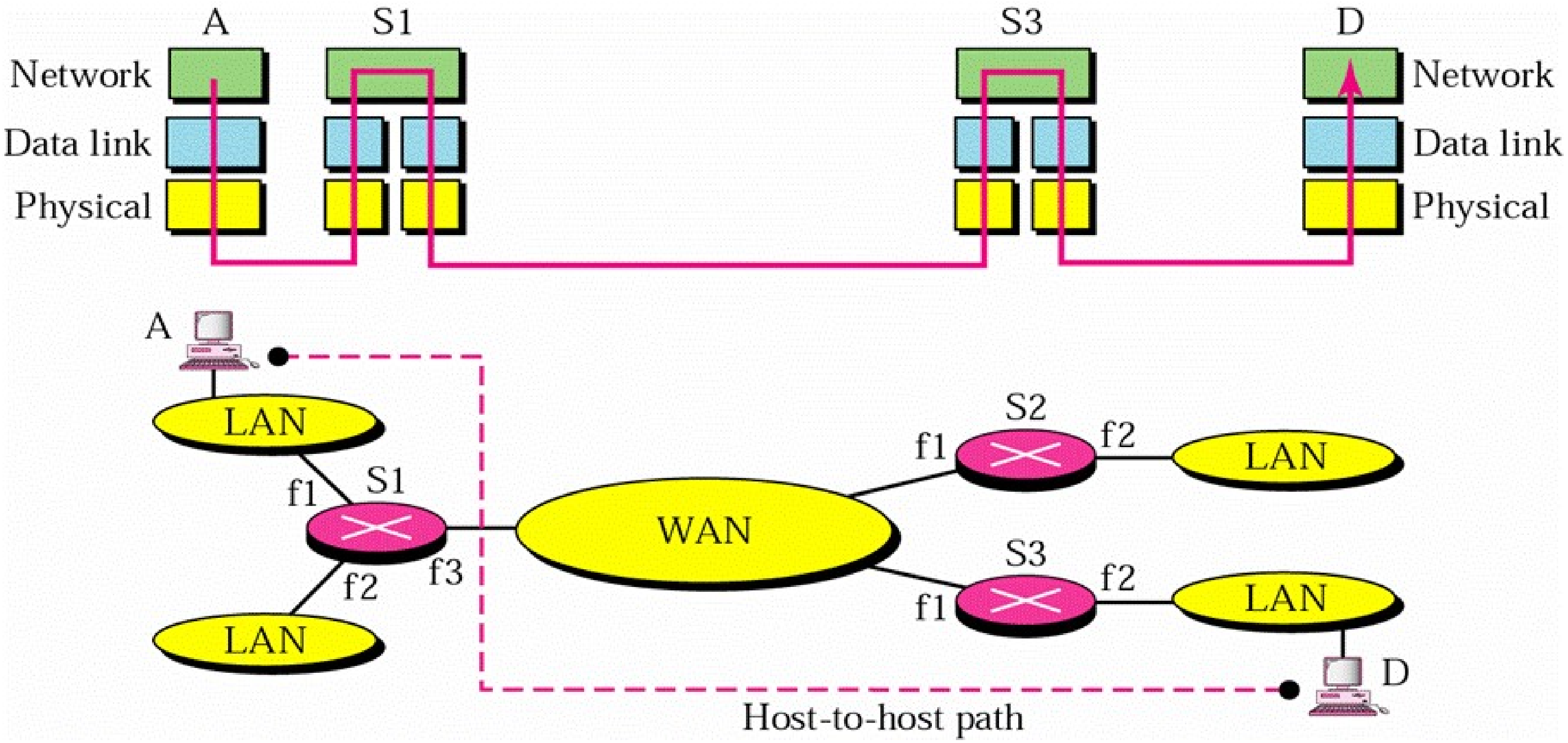
1 hop: adjacent host-router, router-router

Data Link Layer in an Internetwork: Hop-to-Hop Delivery(逐跳传输)



[Forouzan]

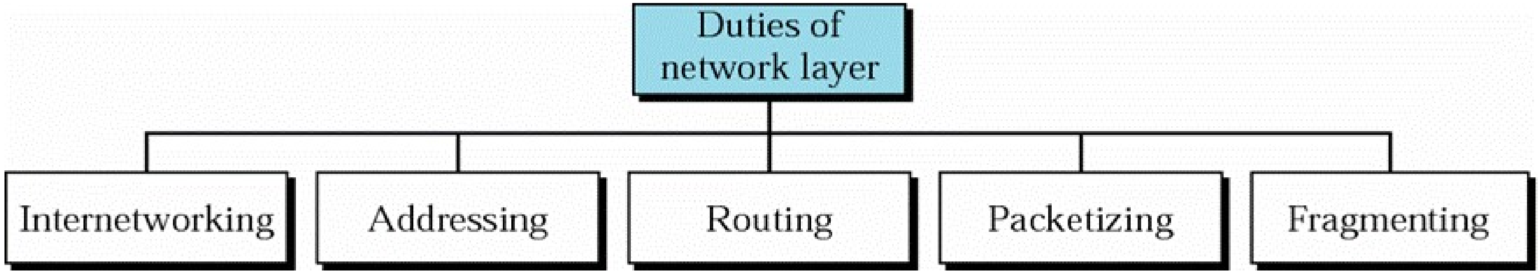
Network Layer in an Internetwork: Host-to-Host Delivery(主机-主机传输)



主机-主机（端-端）

[Forouzan]

Network Duties: Routing



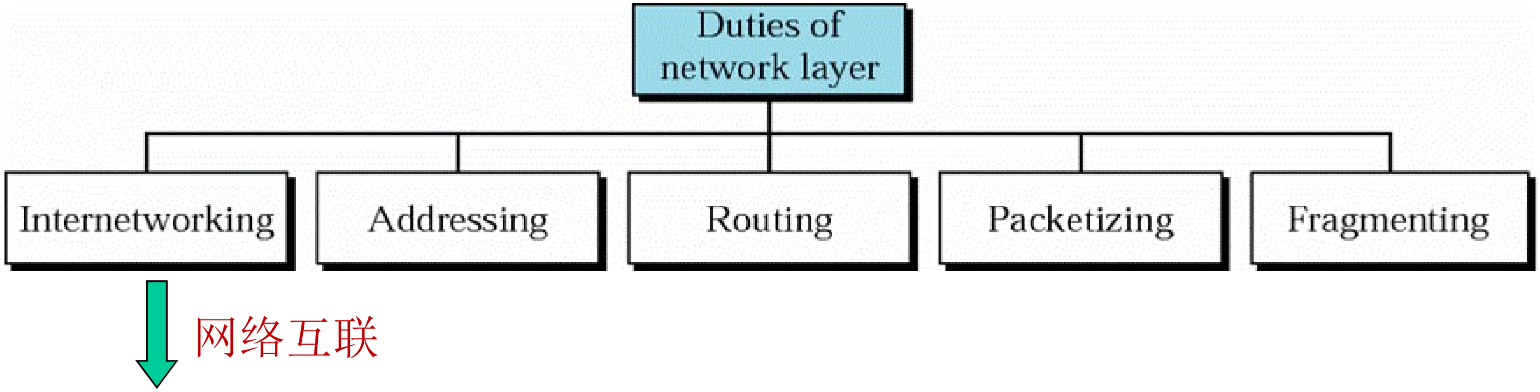
路由选择

To send the packet to destination, who is the **next hop**（下一跳）?

Packet cannot choose the route; the routers connecting the LANs/WANs makes this decision.

[Forouzan]

Network Duties: Internetworking



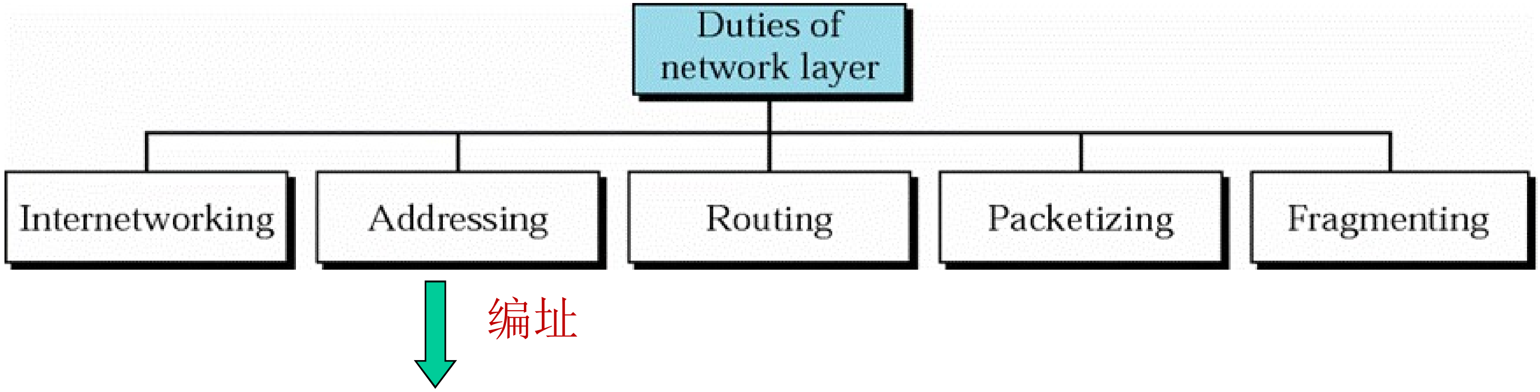
Interconnecting different networks and making them look same to the transport layer.

不同的网络：网络地址不同、实现技术可能不同、包格式可能不同

The transport layer should not be worried about the underlying physical network！（网络的个数及其具体技术对于传输层透明）

[Forouzan]

Network Duties: Addressing

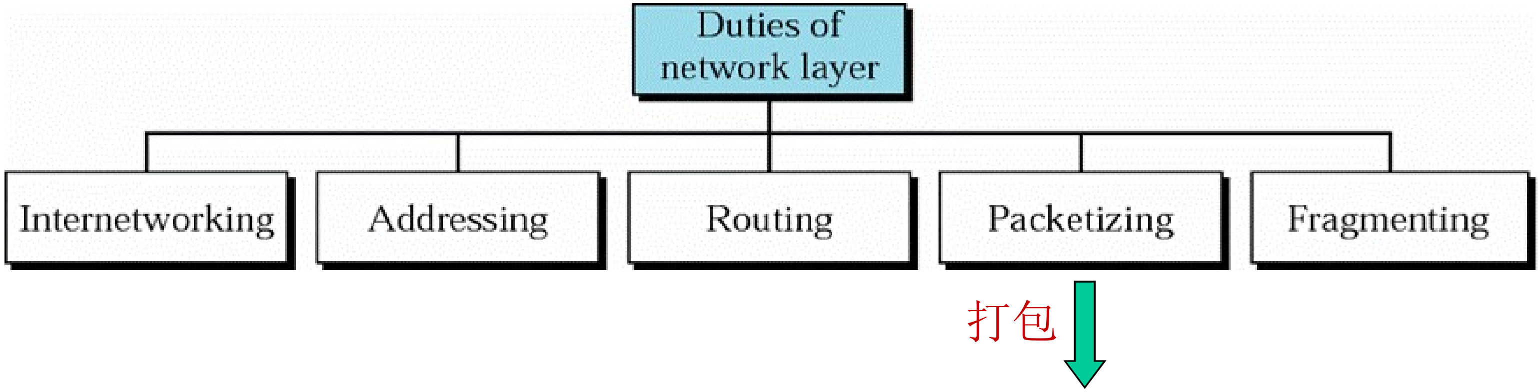


The addresses of a host/router must be uniquely and universally.
(统一的地址格式)

Two devices on the internet can never have the same address.
(Address per interface to network，接口地址唯一)

[Forouzan]

Network Duties: Packetizing

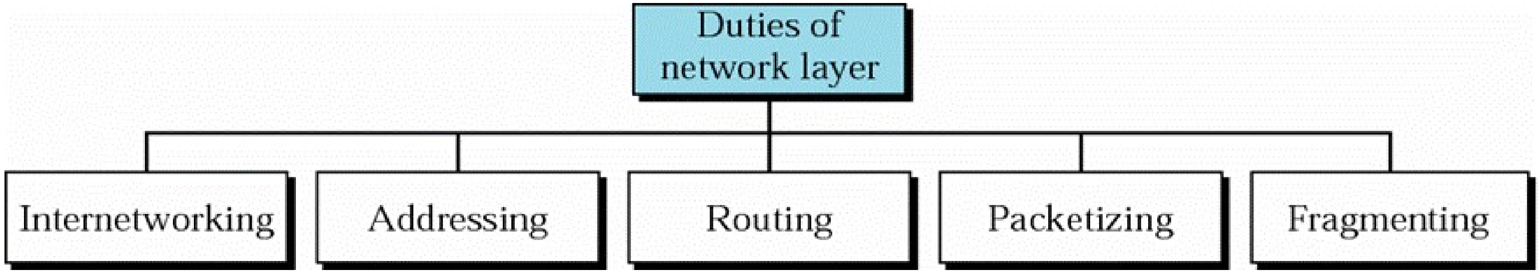


Encapsulated message received from upper layer and forming packets, a task common to all layers.

In Internet, packetizing is done by IP.

[Forouzan]

Network Duties: Fragmenting



分段/分片

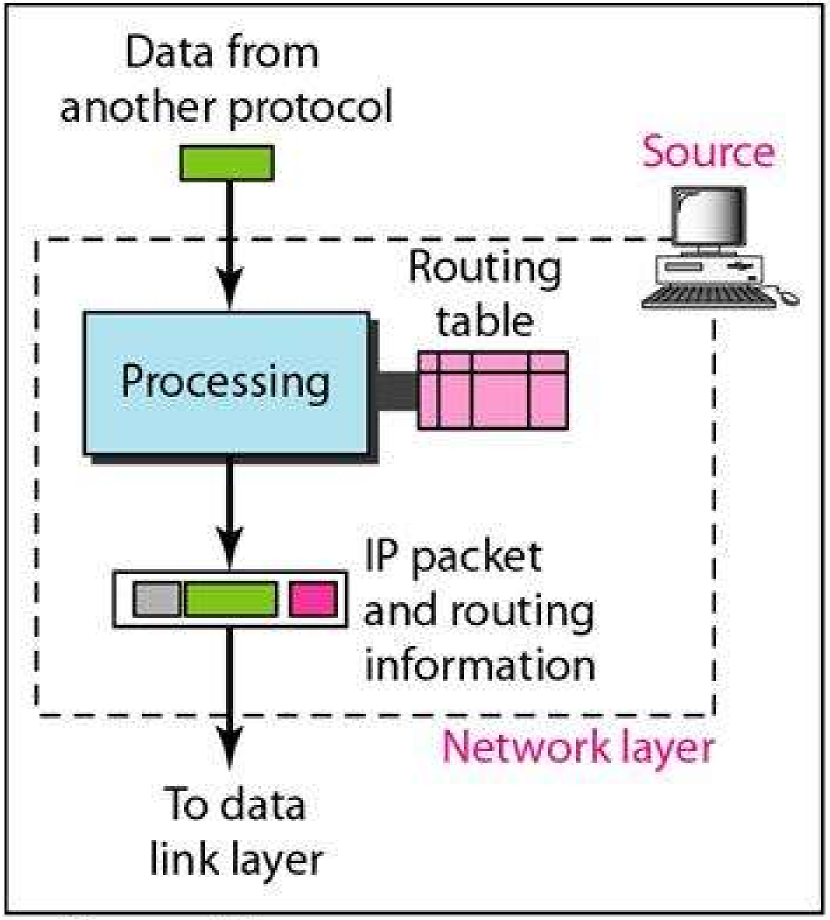
The network layer must be able to operate on top of any data-link layer technology.

All these technologies can handle a different packet length.
（不同的数据链路层网络支持的最大数据长度不同）

The network layer must be able to fragment transport layer PDUs into smaller units so that they can be transferred over various data-link layer technologies.

[Forouzan]

Network Layer at the Source



The source network layer create a packet carrying addresses of source host and destination host.

The source also checks routing table for find the routing information (e.g. outgoing interface, 输出接口, such as Ethernet or wifi).

[Forouzan]

Example of Routing Table in Source

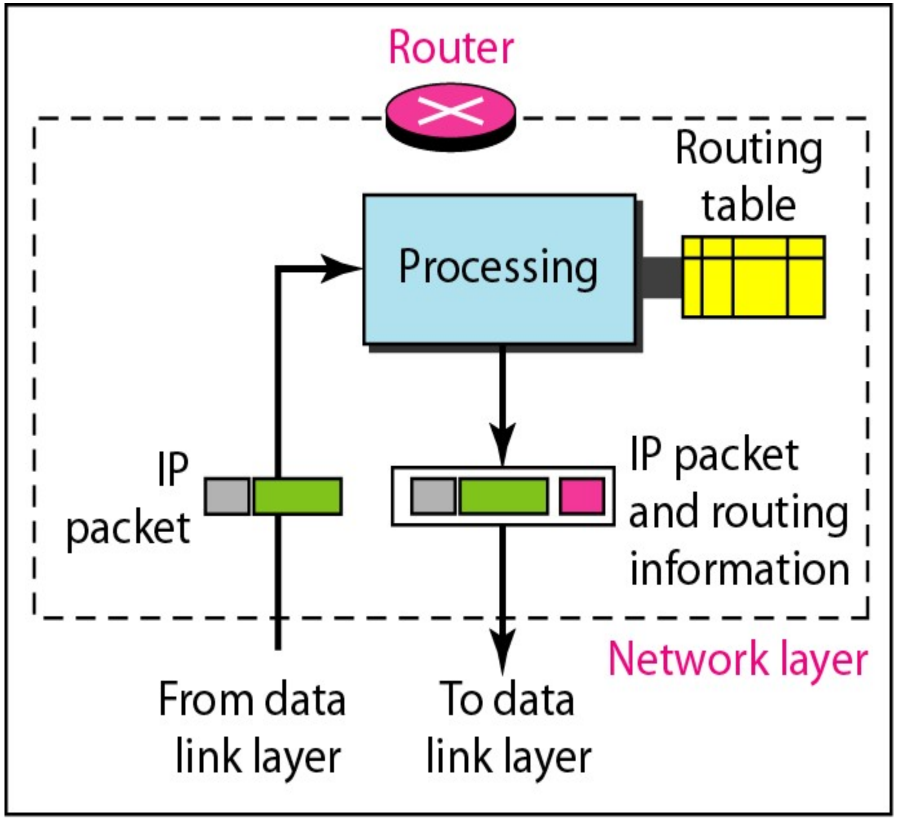
本机IP地址：10.122.230.66
默认网关(路由器)：10.122.192.1

```
C:\Users\Administrator>route print
===========================================================================
接口列表
 10...00 ff 5d 41 0e 4e ......Sangfor aTrust VNIC
 12...54 e1 ad 4c b6 8e ......Intel(R) Ethernet Connection (4) I219-V
 14...cc 2f 71 a7 ec d5 ......Microsoft Wi-Fi Direct Virtual Adapter
  8...cc 2f 71 a7 ec d4 ......Intel(R) Dual Band Wireless-AC 8265
 23...cc 2f 71 a7 ec d8 ......Bluetooth Device (Personal Area Network)
  1...........................Software Loopback Interface 1
===========================================================================

IPv4 路由表
===========================================================================
活动路由:
网络目标          网络掩码            网关          接口     跃点数
          0.0.0.0          0.0.0.0     10.122.192.1    10.122.230.66     35
          2.0.0.0      255.255.255.0            在链路上          2.0.0.1    257
          2.0.0.1  255.255.255.255            在链路上          2.0.0.1    257
        2.0.0.255  255.255.255.255            在链路上          2.0.0.1    257
     10.122.192.0    255.255.192.0            在链路上    10.122.230.66    291
    10.122.230.66  255.255.255.255            在链路上    10.122.230.66    291
   10.122.255.255  255.255.255.255            在链路上    10.122.230.66    291
        127.0.0.0        255.0.0.0            在链路上        127.0.0.1    331
        127.0.0.1  255.255.255.255            在链路上        127.0.0.1    331
  127.255.255.255  255.255.255.255            在链路上        127.0.0.1    331
        224.0.0.0        240.0.0.0            在链路上        127.0.0.1    331
        224.0.0.0        240.0.0.0            在链路上    10.122.230.66    291
  255.255.255.255  255.255.255.255            在链路上        127.0.0.1    331
  255.255.255.255  255.255.255.255            在链路上    10.122.230.66    291
===========================================================================
永久路由:
  无
```

Network Layer at the Router

Responsible for routing the packet



c. Network layer at a router

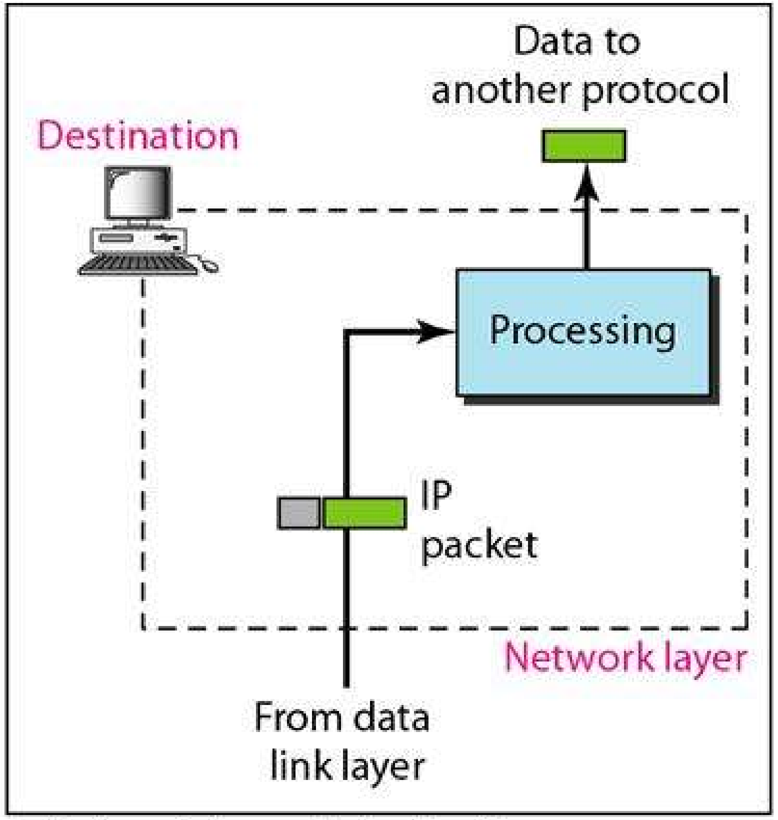
When a packet arrives, the router checks its routing table and finds the interface from which the packet must be sent using routing table.

确定输出接口 → 下一跳

[Forouzan]

Network Layer at the Destination

Responsible for address verification



b. Network layer at destination

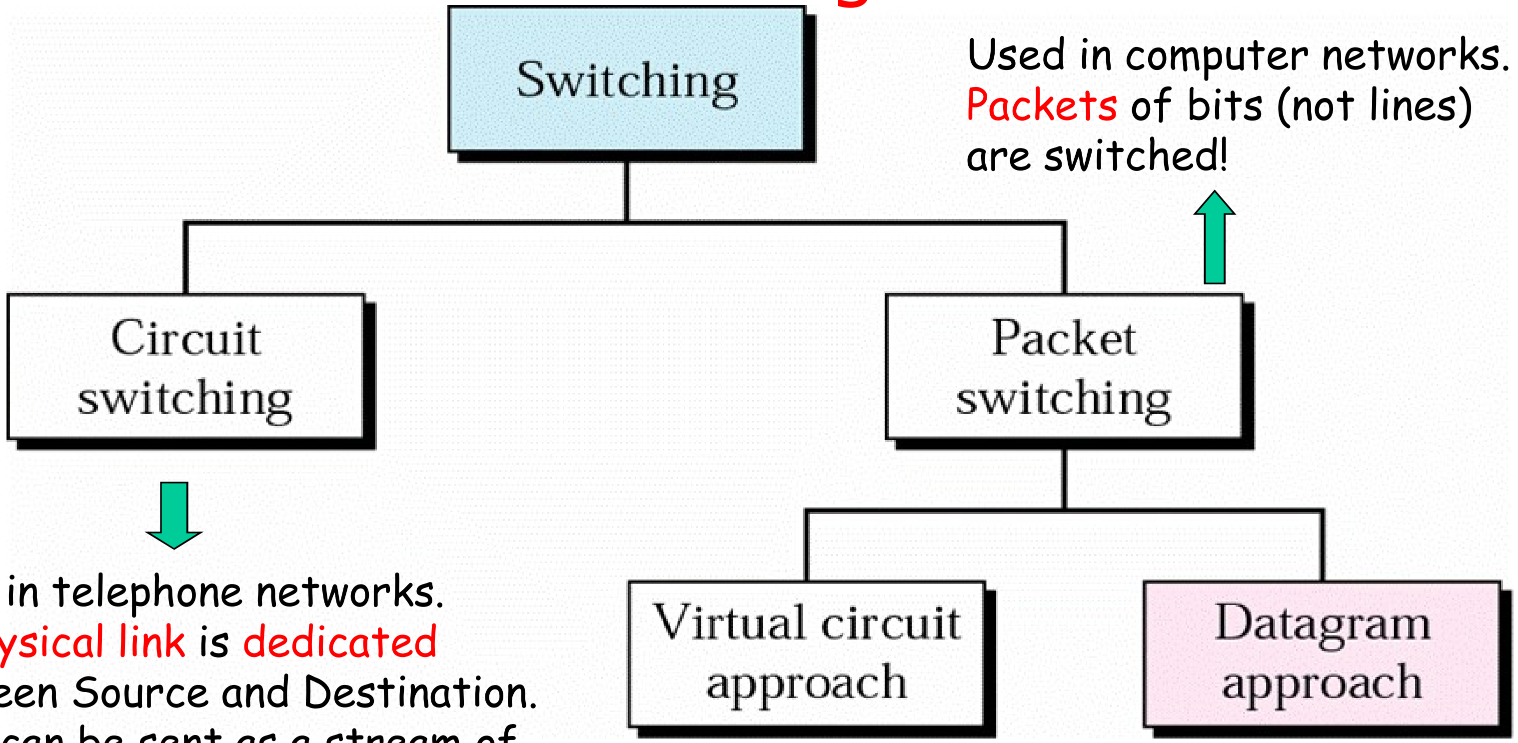
The destination makes sure that the destination address on the packet is the same as the address of the host.

[Forouzan]

Services Provided to Transport Layer

- Design Goal
  - The services should be independent of the router technology.
  - The transport layer should be shielded from the number, type, and topology of the routers present
  - The network addresses made available to the transport layer should use a uniform numbering plan, even across LANs and WANs
- Two classes of service
  - Connection-Oriented Service
  - Connectionless Service

Switching



Used in computer networks. Packets of bits (not lines) are switched!

Used in telephone networks. A physical link is dedicated between Source and Destination. Data can be sent as a stream of bits without the need for packetizing

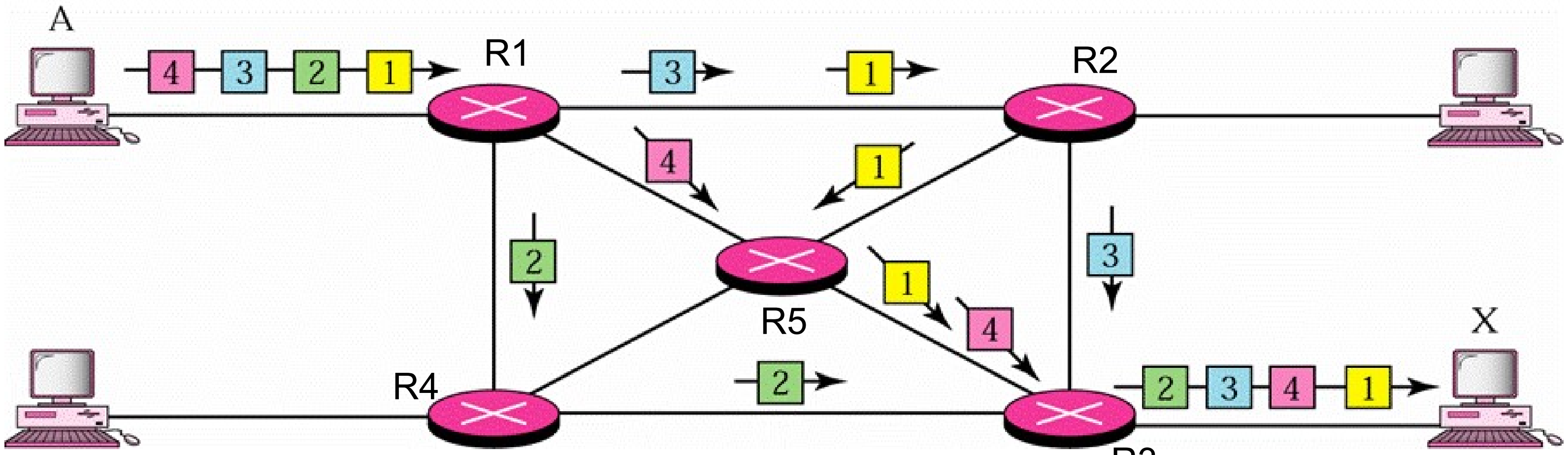
虚电路

Connection-oriented networking

数据报

Connectionless networking

Implementation of Connectionless Service: Datagram(数据报) Network



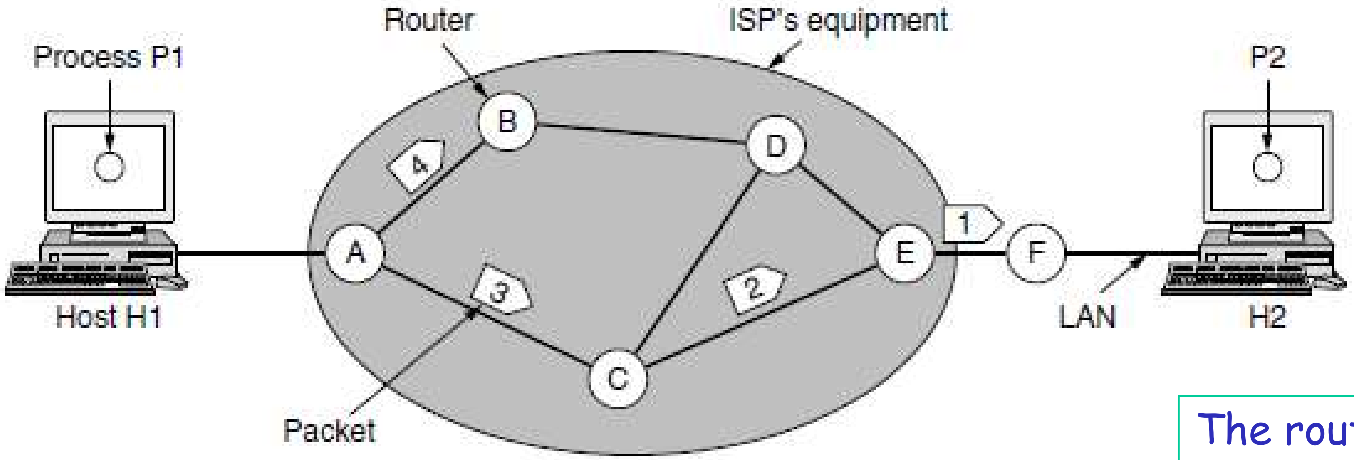
Routing: where to forward the packet?

Routing Table of R1

Dest.	Line(Next Hop，下一跳)
R1	---
R2	R2
R3	R5
R4	R4
R5	R5

[Forouzan]

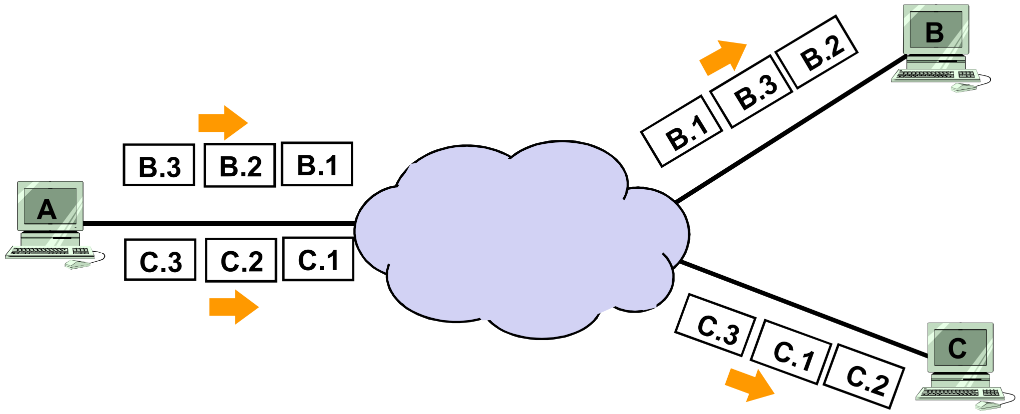
Routing within a Datagram Subnet

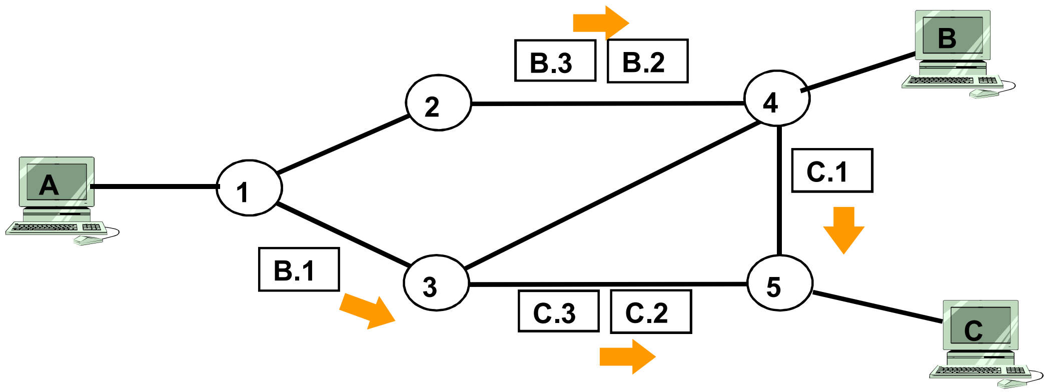


Dest.	A's table (initially) Line	A's table (later) Line	C's table Line	E's table Line
A	–	–	A	C
B	B	B	A	D
C	C	C	–	C
D	B	B	E	D
E	C	B	E	–
F	C	B	E	F

The routing tables are dynamic and are updated periodically.

Routing within a Datagram Network(2)



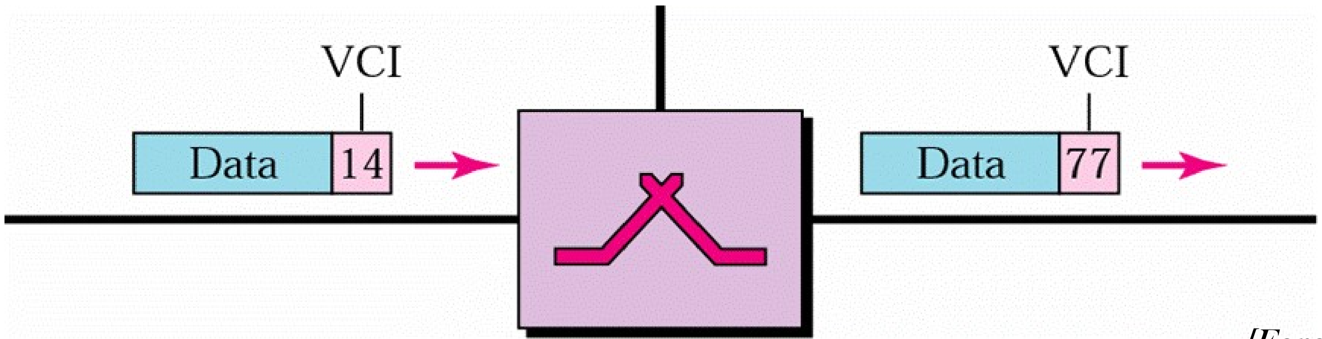


数据报网络实例：
Internet

包中需要携带完整地址；包到达可能乱序

Implementation of Connection-oriented Service: Virtual Circuit(虚电路) Network

- A cross between a circuit-switched network and a datagram network.
- Connection-oriented service
- Virtual circuit is established and end-to-end path is decided before data transmission
- Virtual Circuit Identifier(VCI): local significant
  - Much shorter than address



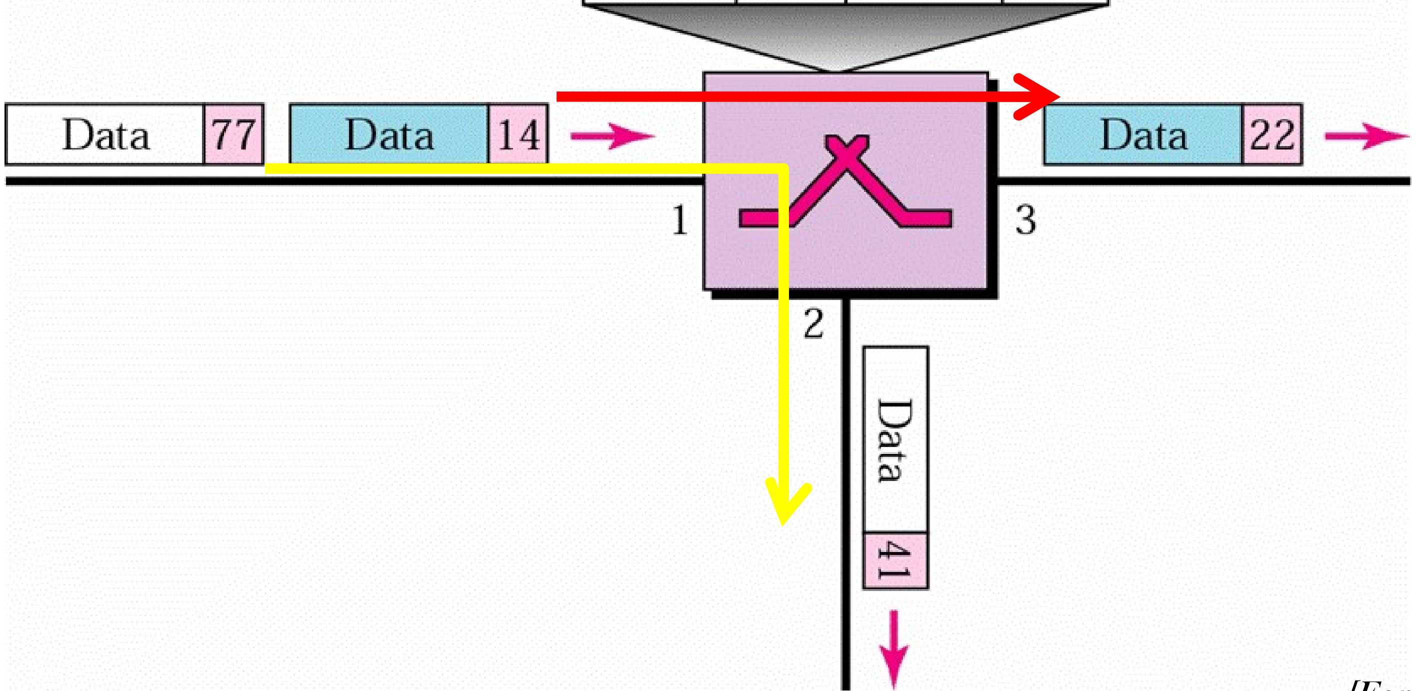
[Forouzan]

Review:

- 网络层的主要功能
  - 端到端（主机-主机）的数据传输
  - 路由选择（选路到下一跳）、网络互联、分段/重装…
- 源主机的网络层：打包、路由选择（确定目的主机在本网内 or 网外）
- 路由器的网络层：路由选择、网络互联
- 目的主机的网络层：核查目的地址…
- 网络层的类型和服务
  - 数据报网络：提供无连接服务，每个包携带完整地址、单独选路、可能乱序
  - 虚电路网络：提供面向连接服务，预先建立连接（选路），包携带较短的 VCI，按序到达
    - VCI用于区分同一个端口/接口上的不同连接

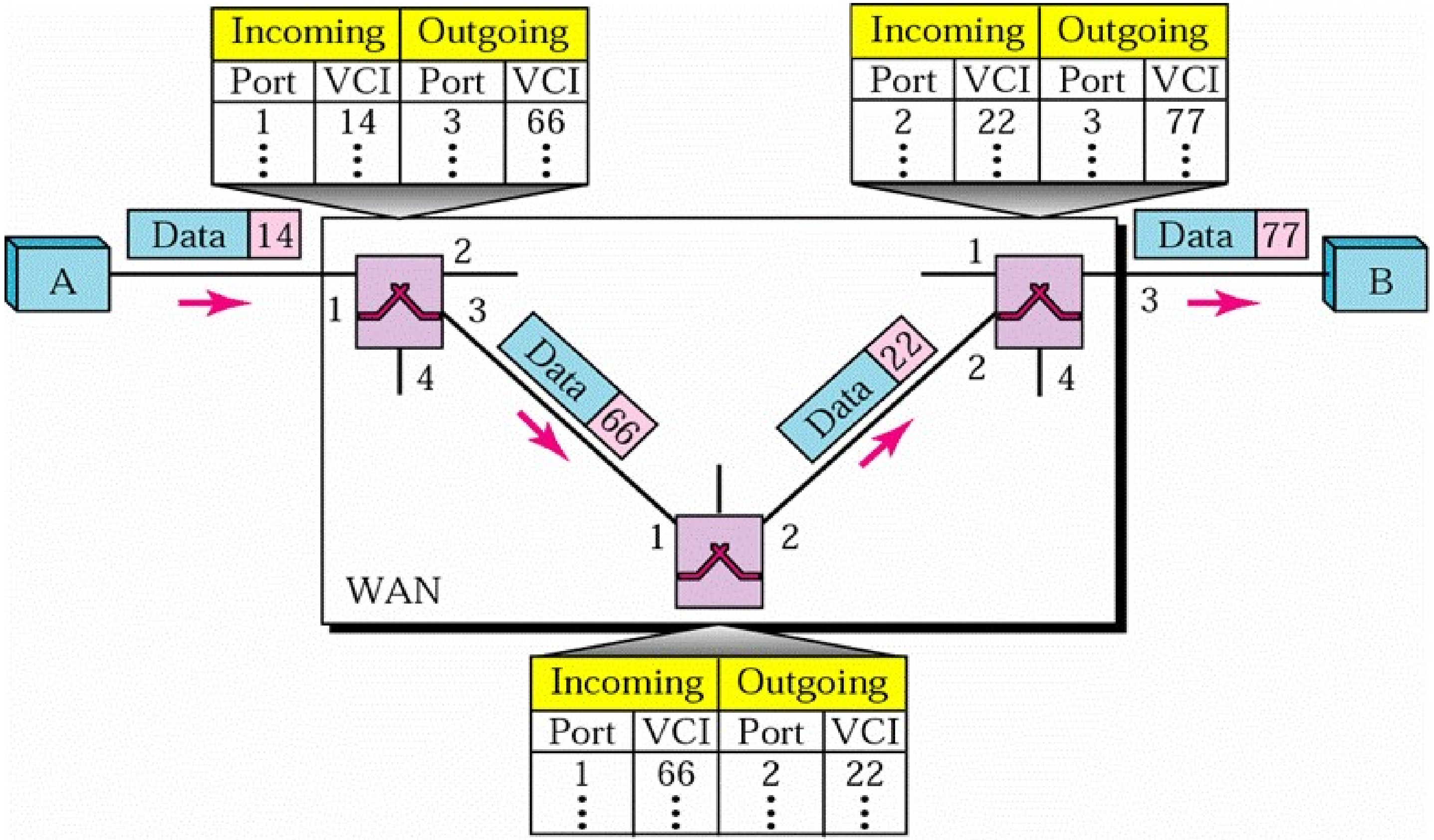
Virtual Circuit Approach: Routing Table

Incoming		Outgoing	
Port	VCI	Port	VCI
1	14	3	22
1	77	2	41



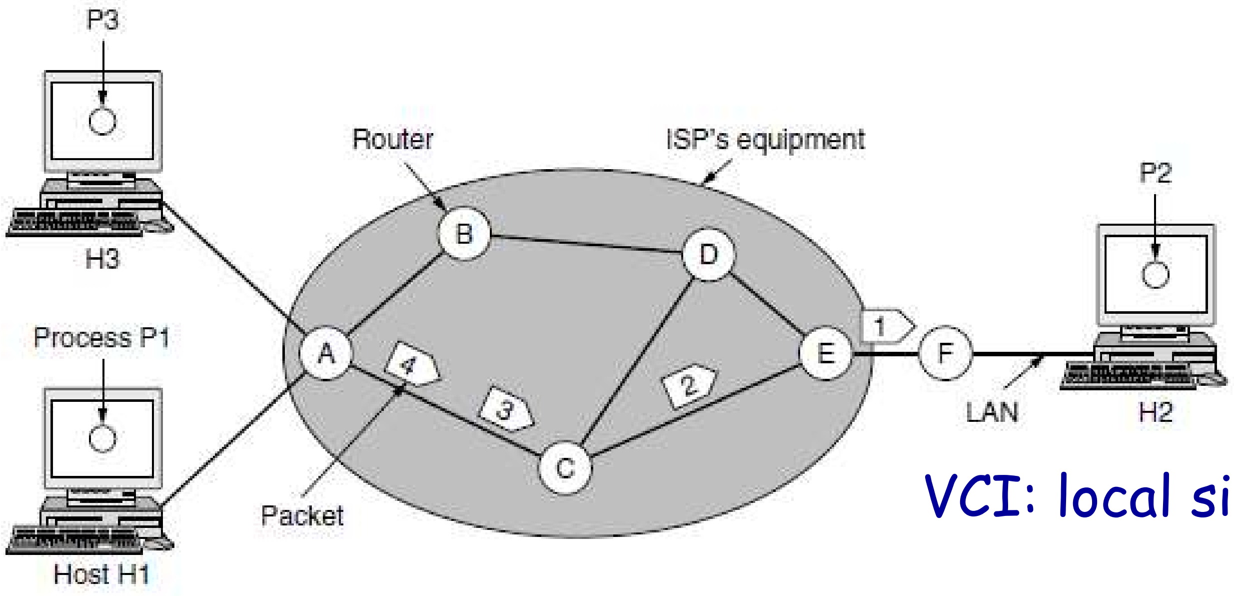
[Forouzan]

Virtual Circuit Approach: Routing



[Forouzan]

Routing within a Virtual-circuit Subnet



VCI: local significant

A's table

In		Out	
H1	1	C	1
H3	1	C	2

C's table

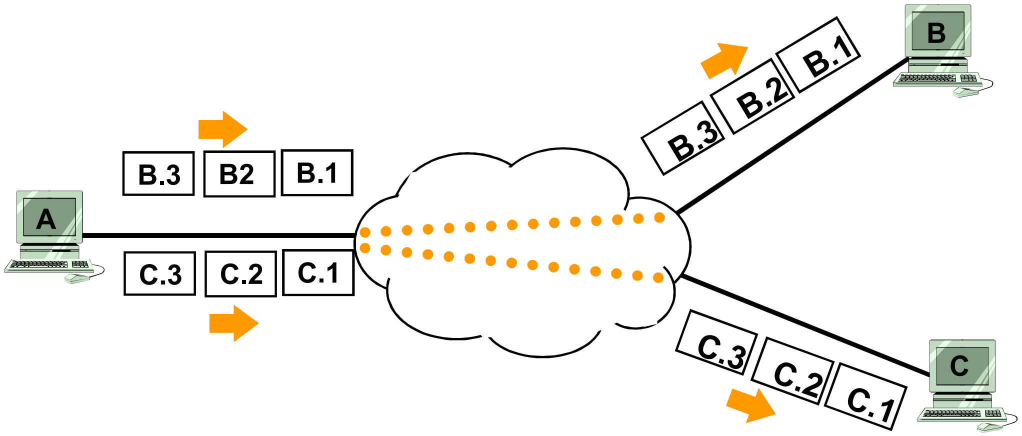
A	1	E	1
A	2	E	2

E's table

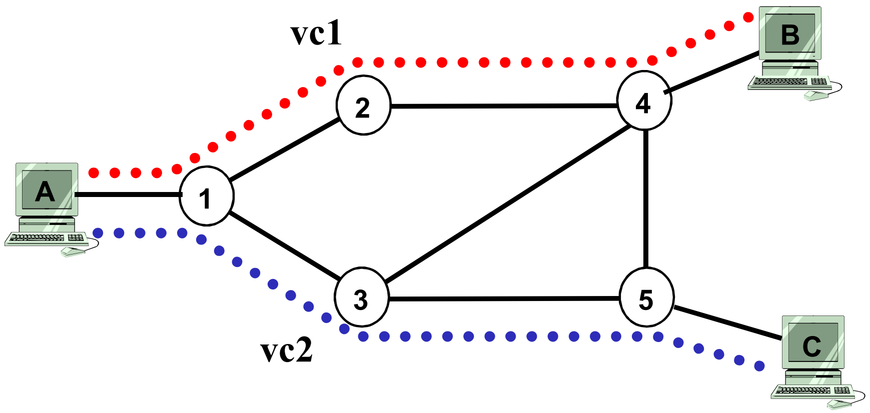
C	1	F	1
C	2	F	2

VC1: H1-A-C-E-F-H2
VC2: H3-A-C-E-F-H2

Routing within a Virtual-circuit Subnet(2)



按序到达



- vc1: A--1--2--4--B
- vc2: A--1--3--5--C

虚电路网络实例：
X.25/帧中继/ATM

Virtual Circuit vs. Datagram

Issue	Datagram network	Virtual-circuit network
Circuit setup	Not needed	Required
Addressing	Each packet contains the full source and destination address	Each packet contains a short VC number
State information	Routers do not hold state information about connections	Each VC requires router table space per connection
Routing	Each packet is routed independently	Route chosen when VC is set up; all packets follow it
Effect of router failures	None, except for packets lost during the crash	All VCs that passed through the failed router are terminated
Quality of service 服务质量	Difficult	Easy if enough resources can be allocated in advance for each VC
Congestion control 拥塞控制	Difficult	Easy if enough resources can be allocated in advance for each VC

Outline

- Introduction to Network Layer
- *Routing Algorithms*
  - *Concept of Routing and Routing Table*
  - Optimality principle
  - Flooding
  - Dijkstra's "Shortest Path" Algorithm
  - Distance Vector Routing (DVR)
  - Link State Routing (LSR)
  - Hierarchical Routing
- Congestion Control and Internetworking
- The network layer in Internet

Routing: Overview(1)

- An internet(互联网) is a combination of networks connected by routers.
- When a packet goes from a source to a destination, it may pass through many routers.
- A router consults a routing table when a packet is ready to be forwarded. The routing table specifies the optimum path (最佳路径) for the packet.
- To make the lookup efficiency, need to make the size of table manageable and handles issues such as security at the same time.
    - Key question is how to design the routing table.

Routing: Overview(2)

- Routing table can be Static or Dynamic.
  - Static table does not change frequently.

> A static routing table
> contains information entered manually
> Usually remained unchanged.

  - Dynamic table is updated automatically when there is a change somewhere in the network.

> A dynamic routing table is updated
> periodically or whenever necessarily
> using one of the dynamic routing protocols
> such as RIP, OSPF, or BGP.

Routing: Overview(3)

- Routing Algorithms: algorithms used to construct and update routing tables
  - DVR, LSR, PVR

- Routing Protocols: a combination of rules/procedures that lets routers inform one another when changes occur; mostly based on sharing/combining information between routers at different networks.
  - RIP, OSPF, BGP

Routing table: Next-hop(下一跳) routing

Routing table for host A

Destination	Route
Host B	R1, R2, Host B

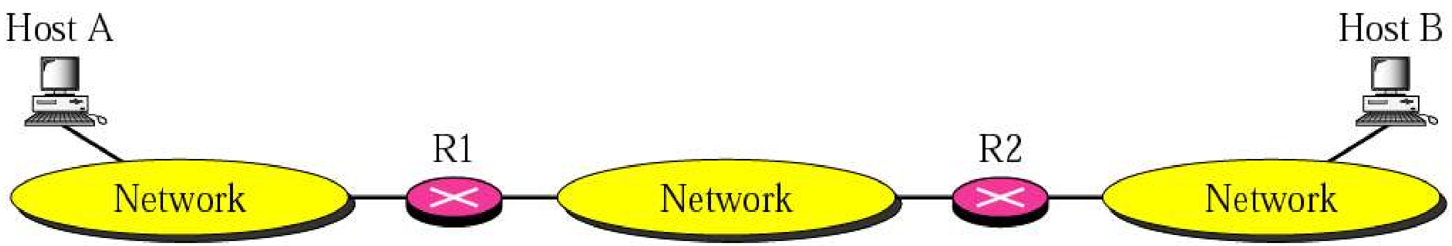
Routing table for R1

Destination	Route
Host B	R2, Host B

Routing table for R2

Destination	Route
Host B	Host B

a. Routing tables based on route （完整路径）



Routing table for host A

Destination	Next Hop
Host B	R1

Routing table for R1

Destination	Next Hop
Host B	R2

Routing table for R2

Destination	Next Hop
Host B	—

b. Routing tables based on next hop (下一跳选路)

[Forouzan]

Routing Algorithms

- Concept of Routing and Routing Table
- *Optimality principle*
- Flooding
- Dijkstra's "Shortest Path" Algorithm
- Distance Vector Routing (DVR)
- Link State Routing (LSR)
- Hierarchical Routing

How to evaluate routing algorithms?

- Goals
  - Correctness
  - Simplicity
  - Robustness 健壮
  - Stability 稳定
  - Fairness 公平
  - Optimality 最优
    - Minimizing mean packet delay 时延最小
    - Maximizing total network throughput 吞吐量最大
- Performance Criteria
  - Least cost, such as minimum hops

Responsibility of Routing Algorithms

- Responsible for deciding which output line an incoming packet should be transmitted on
  - If the subnet uses datagrams internally, this decision must be made for every arriving data packet
    数据报网络对每个包选路
  - If the subnet uses virtual circuits internally, routing decisions are made only when a new virtual circuit is being set up. Thereafter, data packets just follow the previously-established route.
    虚电路网络，只在建立虚电路时选路
- Routing(选路) & Forwarding(转发)

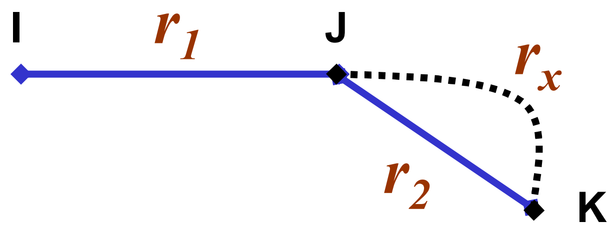
Classification of Routing Algorithms

- *Nonadaptive* algorithms (Static Routing)
  - The choice of the route to use is computed in advance, off-line, and downloaded to the routers when the network is booted
  - The routing decisions are not based on measurements or estimates of the current traffic and topology

- *Adaptive(自适应)* algorithms
  - Changing their routing decisions to reflect changes in the topology, and usually the traffic as well
    适应网络拓扑和业务量的变化

The Optimality Principle(最优化原则)

- If router **J** is on the optimal path from router **I** to router **K**, then the optimal path from **J** to **K** also falls along the same route
  - If a route r_x better than r_2 existed from **J** to **K**, it could be concatenated with r_1 to improve the route from **I** to **K**
  - Contradicting our statement that r_1r_2 is optimal



r_1+r_2

r_1+r_x

Major Routing Strategies

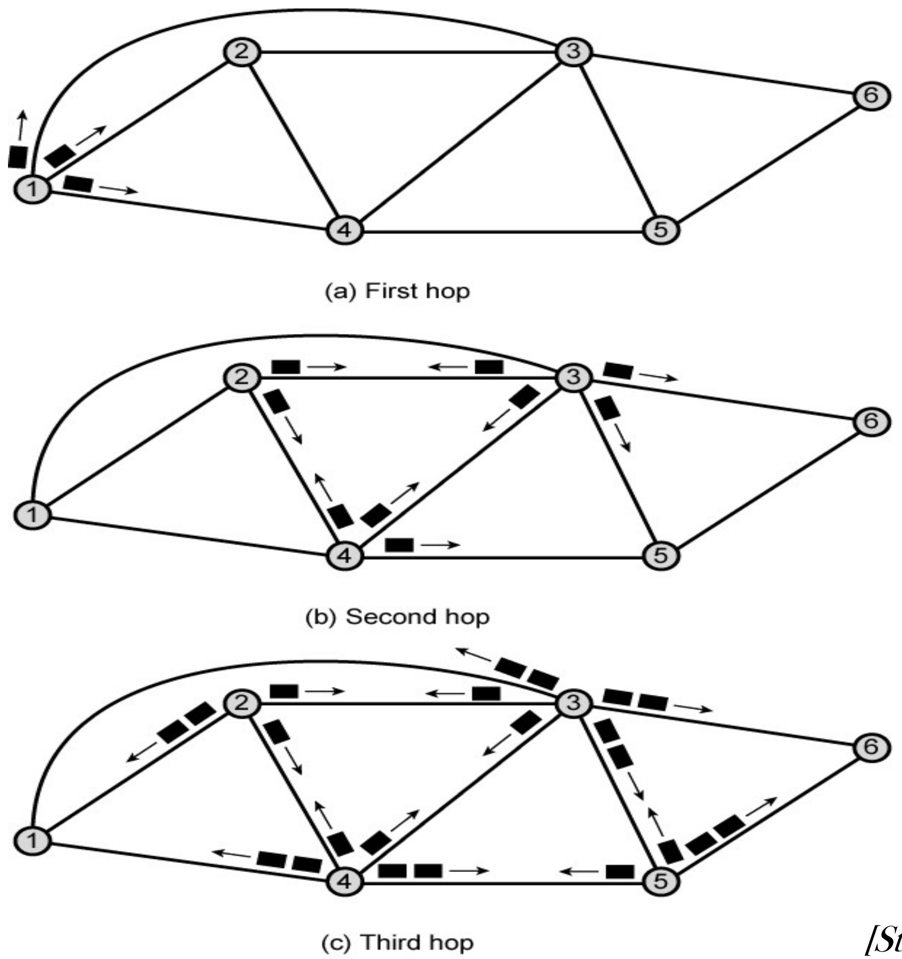
- Fixed Routing(固定查表选路)
  - Single permanent route for each source to destination pair
- Flooding(洪泛/泛洪)
  - Packet sent to every neighbor
  - Incoming packets forwarded to every link except incoming link
- Random Routing(随机选路)
  - Selects one outgoing path to forward incoming packet
  - Selection can be random or round robin
- Adaptive Routing(自适应选路)
  - Routing decisions change as conditions on the network change
  - Requires information about network
  - Used by almost all packet switching networks

Outline

- Introduction to Network Layer
- *Routing Algorithms*
  - Concept of Routing and Routing Table
  - Optimality principle
  - *Flooding*
  - Dijkstra's "Shortest Path" Algorithm
  - Distance Vector Routing (DVR)
  - Link State Routing (LSR)
  - Hierarchical Routing
- Congestion Control and Internetworking
- The network layer in Internet

Flooding(洪泛法): Ideas

- 无需路由表
- No network information required
- Incoming packets retransmitted on every link except incoming link
- Eventually a number of copies will arrive at destination



[Stallings]

Figure 12.4 Flooding Example (hop count = 3)

Flooding: Problems and Solutions

- Problem
  - Generates vast numbers of duplicate packets
- Solutions
  - The header of each packet contains a hop counter(跳计数器)
    - Counter is decremented at each hop
    - When the counter reaches zero, discard the packet
    - Sender initializes the hop counter as the length of the path from source to destination, or the subnet diameter
  - Source puts a sequence number in each packet（序列号）
    - Each router records maximum seq per source telling which seq have already been seen

Flooding: Properties

- All possible routes are tried
  - Very robust 健壮
- At least one packet will have taken minimum hop count route (shortest path)
- All nodes are visited
  - Useful to distribute routing information

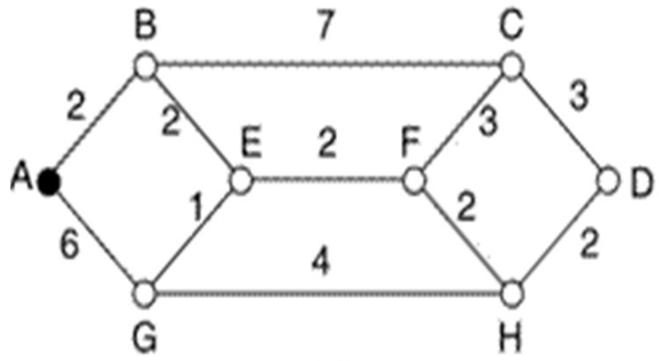
<u>*Survey:*</u> where are flooding used?

Outline

- Introduction to Network Layer
- *Routing Algorithms*
  - Concept of Routing and Routing Table
  - Optimality principle
  - Flooding
  - *Dijkstra's "Shortest Path" Algorithm*
  - Distance Vector Routing (DVR)
  - Link State Routing (LSR)
  - Hierarchical Routing
- Congestion Control and Internetworking
- The network layer in Internet

Shortest Path Routing(最短路径选路)

- Build a weighted and directed graph of subnet(有向图)
  - Nodes represent routers
  - Arcs represents communication lines

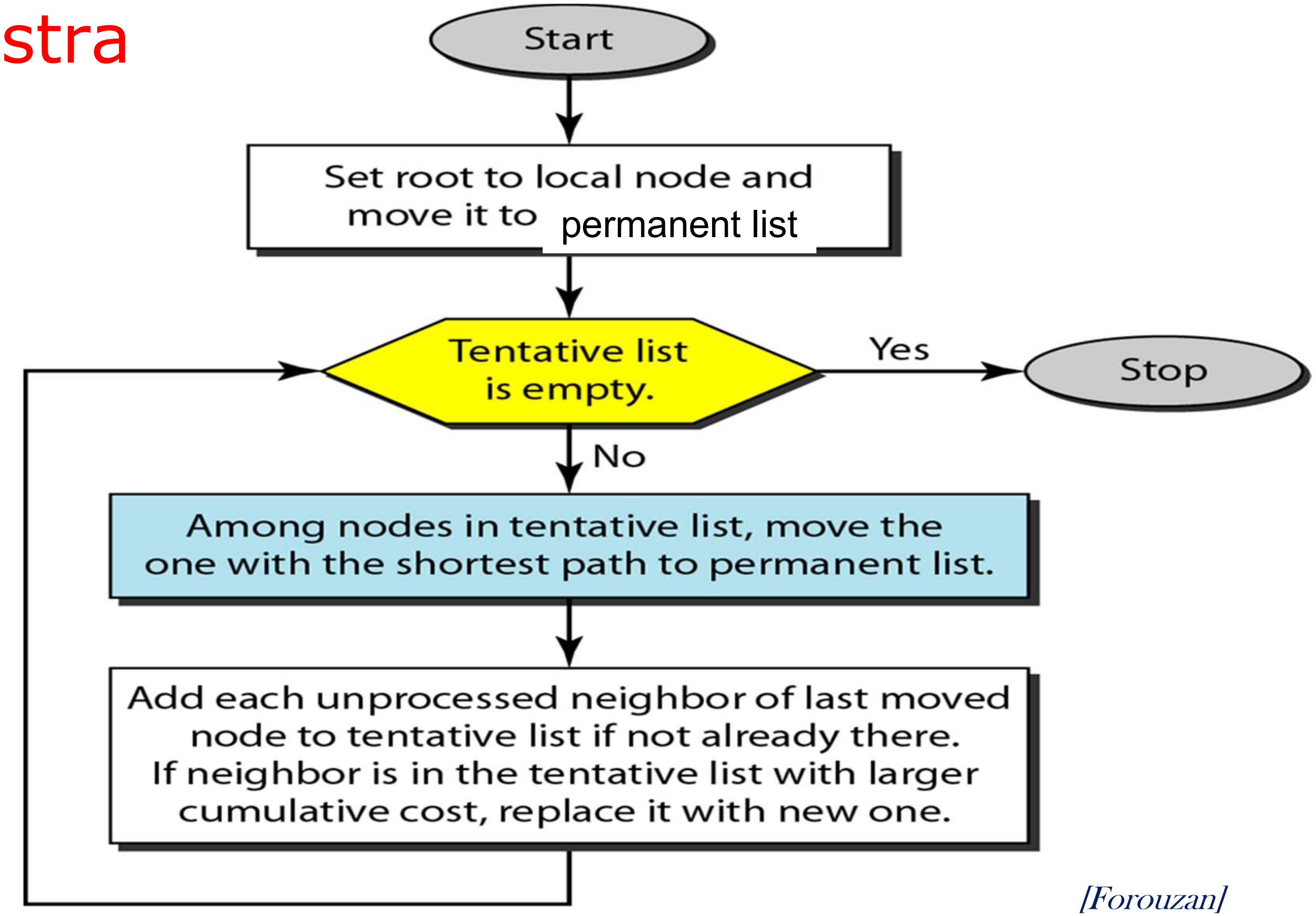


- Weight
  - Hops
  - Geographic distance in kilometers
  - Mean queuing and transmission delay
  - A function of the distance, bandwidth, average traffic, communication cost, mean queue length, measured delay, and other factors
- Dijkstra Algorithm
  - Finds the shortest path between a given pair of routers

Dijkstra Algorithm

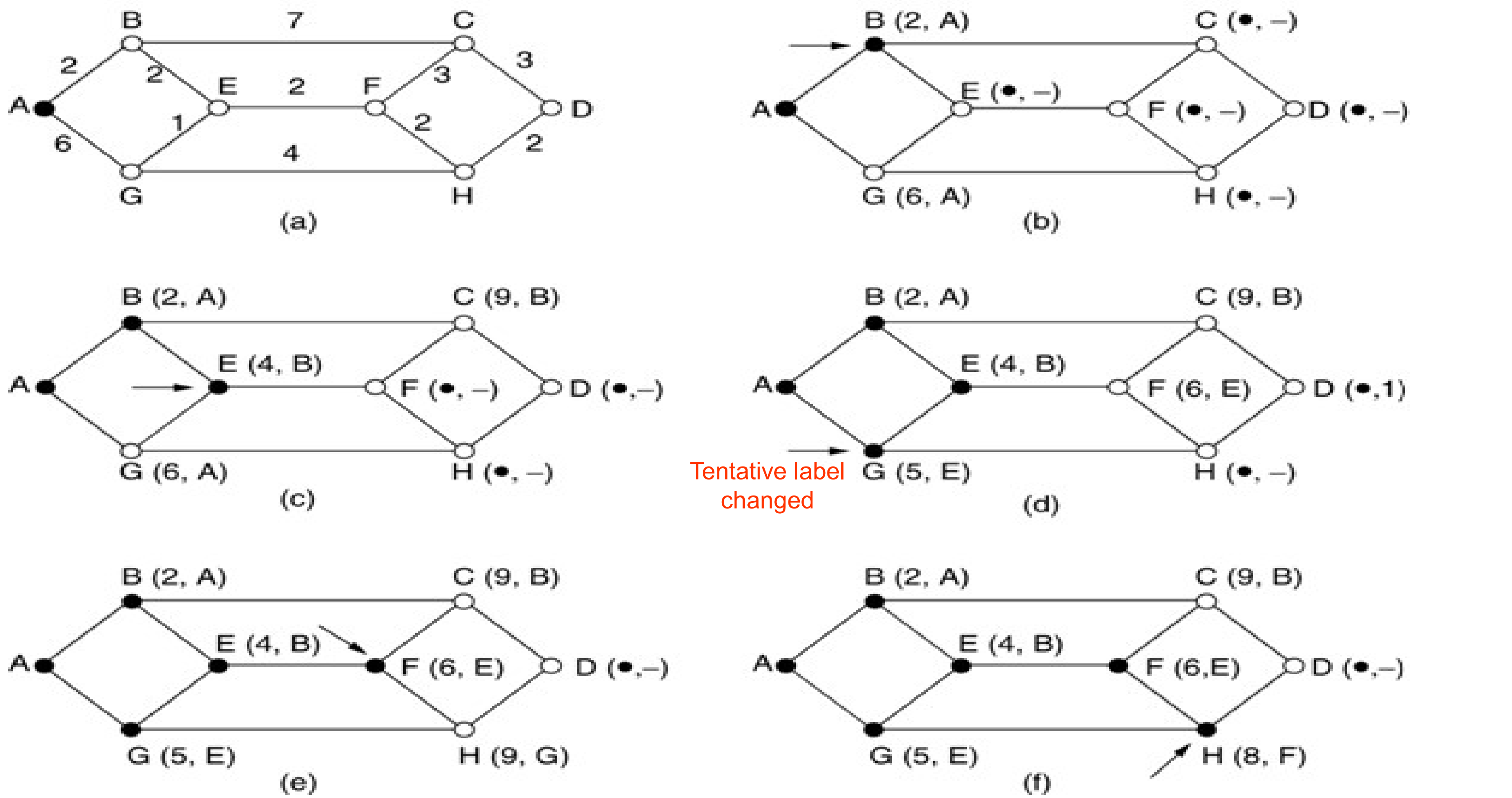
- Find shortest paths from given source node to all other nodes
- Routing Algorithm
  - 1. Starting with the local node (router): the root of the tree.
  - 2. Assign a cost of 0 to this node and make it the first permanent node（永久节点).
  - 3. Examine each neighbor node of the node that was the last permanent node.
  - 4. Assign a cumulative cost（累计开销） to each node and make it tentative（临时节点）.
  - 5. Among the list of tentative nodes
    - Find the node with the smallest cumulative cost and make it permanent.
    - If a node can be reached from more than one direction, Select the direction with the shortest cumulative cost.
  - 6. Repeat steps 3 to 5 until every node becomes permanent.

Flow of Dijkstra

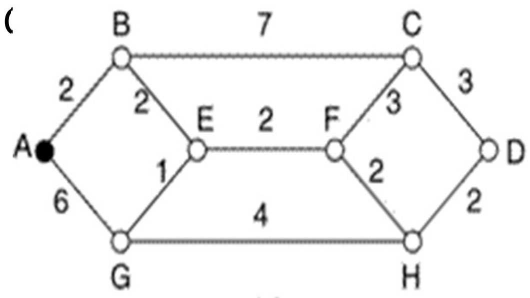


[Forouzan]

Example of Dijkstra Algorithm

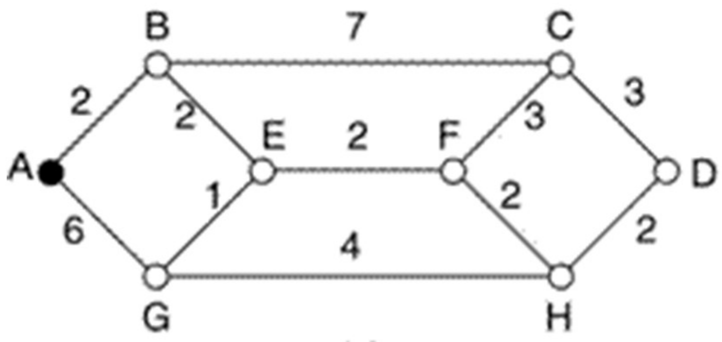


Example of Dijkstra



A	B	C	D	E	F	G	H	Permanent Node
∞	∞	∞	∞	∞	∞	∞	∞	
0	∞	∞	∞	∞	∞	∞	∞	A
0	2, A	∞	∞	∞	∞	6, A	∞	B
0	---	9, B	∞	4, B	∞	6, A	∞	E
0	---	9, B	∞	---	6, E	5, E	∞	G
0	---	9, B	∞	---	6, E	---	9, G	F
0	---	9, B	∞	---	---	---	8, F	H
0	---	9, B	10,H	---	---	---	---	C
0	---	---	10,H	---	---	---	---	D

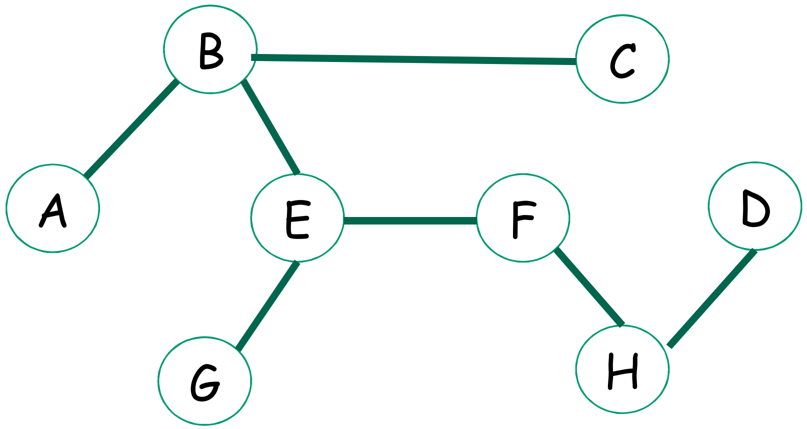
What we get by using Dijkstra?



A's routing table

Dest.	Cost	Next Hop
A	0	--
B	2	B
C	9	B
D	10	B
E	4	B
F	6	B
G	5	B
H	8	B

A's shortest path tree



Outline

- Introduction to Network Layer
- *Routing Algorithms*
  - Concept of Routing and Routing Table
  - Optimality principle
  - Flooding
  - Dijkstra's "Shortest Path" Algorithm
  - *Distance Vector Routing (DVR)*
  - Link State Routing (LSR)
  - Hierarchical Routing
- Congestion Control and Internetworking
- The network layer in Internet

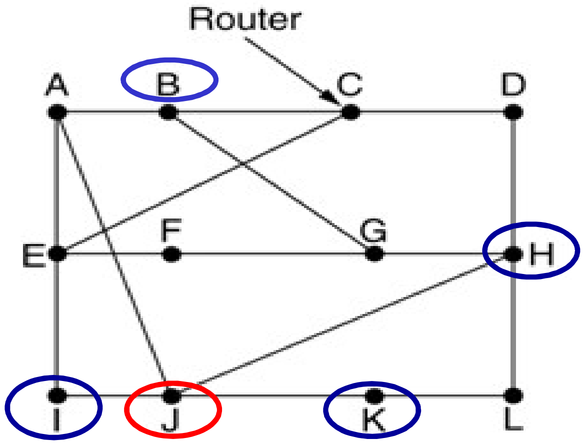
Distance Vector Routing(距离矢量选路)

- Each router maintains a routing table (Vector)
  - The best known distance to each destination(距离)
    - The distance might be number of hops, time delay in milliseconds, …
  - Which line to use (下一跳/接口)
- The table is updated by exchanging information with neighbors
  - Routing information sent: **(Destination, Distance)**
    - Periodically (RIP:30s)
  - When routing information received
    - Router updates local table if receive a better route
    - Refreshes existing routes (old routing table is not used in the calculation.)
  - Delete routing table items if they time out (RIP: 180s)

AB ≠ BA

No next hop info.

DVR: Example in textbook



(a)

To	A	I	H	K	New estimated delay from J	Line
A	0	24	20	21	8	A
B	12	36	31	28	20	A
C	25	18	19	36	28	I
D	40	27	8	24	20	H
E	14	7	30	22	17	I
F	23	20	19	40	30	I
G	18	31	6	31	18	H
H	17	20	0	19	12	H
I	21	0	14	22	10	I
J	9	11	7	10	0	–
K	24	22	22	0	6	K
L	29	33	9	9	15	K
	JA delay is 8	JI delay is 10	JH delay is 12	JK delay is 6	New routing table for J	

Vectors received from J's four neighbors

(b)

+

Taking destination B as an example,

JB= JA+AB= 8+12=20

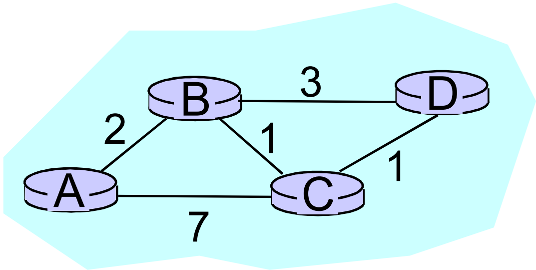
JB= JI+IB = 10+36=46

JB= JH+HB= 12+31=43

JB= JK+KB= 6+28=34

Thus, A is selected as next hop.

DVR Example 2: Initially



Node A

Dest.	Cost	NextHop
B	2	B
C	7	C
D	∞	-

Node B

Dest.	Cost	NextHop
A	2	A
C	1	C
D	3	D

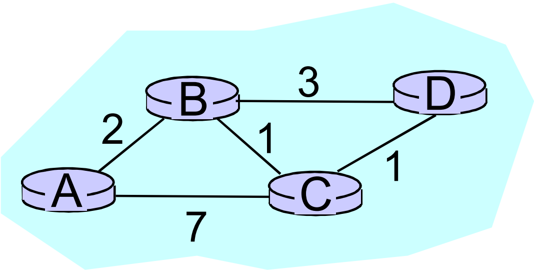
Node C

Dest.	Cost	NextHop
A	7	A
B	1	B
D	1	D

Node D

Dest.	Cost	NextHop
A	∞	-
B	3	B
C	1	C

Example 2: 1st Iteration (C → A)



Node A

Dest.	Cost	NextHop
B	2	B
C	7	C
D	8	C

Node B

Dest.	Cost	NextHop
A	2	A
C	1	C
D	3	D

D(A, D) = D(A, C) + D(C,D) = 7 + 1 = 8

(D(C,A), D(C,B), D(C,D))

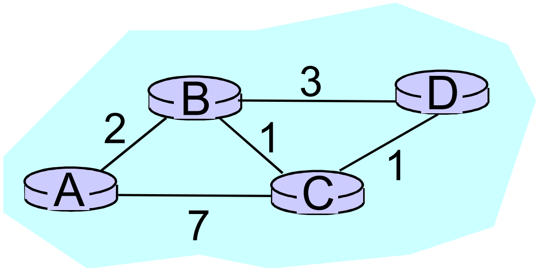
Node C

Dest.	Cost	NextHop
A	7	A
B	1	B
D	1	D

Node D

Dest.	Cost	NextHop
A	∞	-
B	3	B
C	1	C

Example 2: 1st Iteration (B→A)



Node A

Dest.	Cost	NextHop
B	2	B
C	3	B
D	5	B

Node B

Dest.	Cost	NextHop
A	2	A
C	1	C
D	3	D

D(A,D) = D(A,B) + D(B,D) = 2 + 3 = 5

D(A,C) = D(A,B) + D(B,C) = 2 + 1 = 3

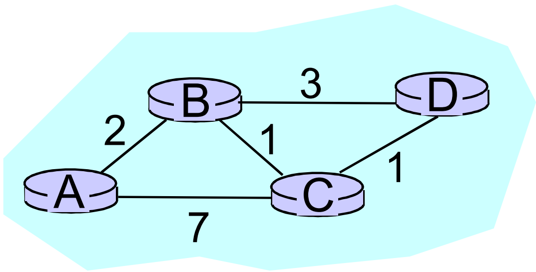
Node C

Dest.	Cost	NextHop
A	7	A
B	1	B
D	1	D

Node D

Dest.	Cost	NextHop
A	∞	-
B	3	B
C	1	C

Example 2: End of 1st Iteration



Node A

Dest.	Cost	NextHop
B	2	B
C	3	B
D	5	B

Node B

Dest.	Cost	NextHop
A	2	A
C	1	C
D	**2**	**C**

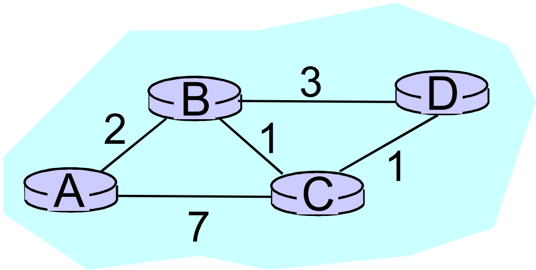
Node C

Dest.	Cost	NextHop
A	**3**	**B**
B	1	B
D	1	D

Node D

Dest.	Cost	NextHop
A	**5**	**B**
B	**2**	**C**
C	1	C

Example 2: End of 2nd Iteration



Node A

Dest.	Cost	NextHop
B	2	B
C	3	B
D	4	B

Node B

Dest.	Cost	NextHop
A	2	A
C	1	C
D	2	C

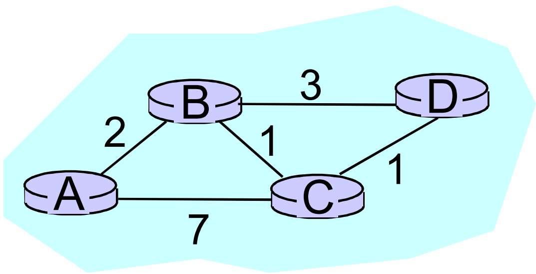
Node C

Dest.	Cost	NextHop
A	3	B
B	1	B
D	1	D

Node D

Dest.	Cost	NextHop
A	4	C
B	2	C
C	1	C

Example 2: End of 3rd Iteration



Node A

Dest.	Cost	NextHop
B	2	B
C	3	B
D	4	B

Node B

Dest.	Cost	NextHop
A	2	A
C	1	C
D	2	C

Node C

Dest.	Cost	NextHop
A	3	B
B	1	B
D	1	D

Node D

Dest.	Cost	NextHop
A	4	C
B	2	C
C	1	C

Nothing changes → algorithm terminates

Features and Problems of DVR

- DVR: Build/update routing table based on information from neighbors
- 3 key features
  - Sharing knowledge about the entire network with neighbors (distance to all nodes)
  - Sharing only with immediate neighbors.
  - Sharing at regular intervals.
- Drawbacks:
  - Slow convergence(收敛慢)
  - Reacts rapidly to good news, but leisurely to bad news

DVR：Count-to-Infinity(无穷计数) Problem

- *Also regarded as "Infinite Loop Problem"*

A	B	C	D	E	
	•	•	•	•	Initially
	1	•	•	•	After 1 exchange
	1	2	•	•	After 2 exchanges
	1	2	3	•	After 3 exchanges
	1	2	3	4	After 4 exchanges

(a)

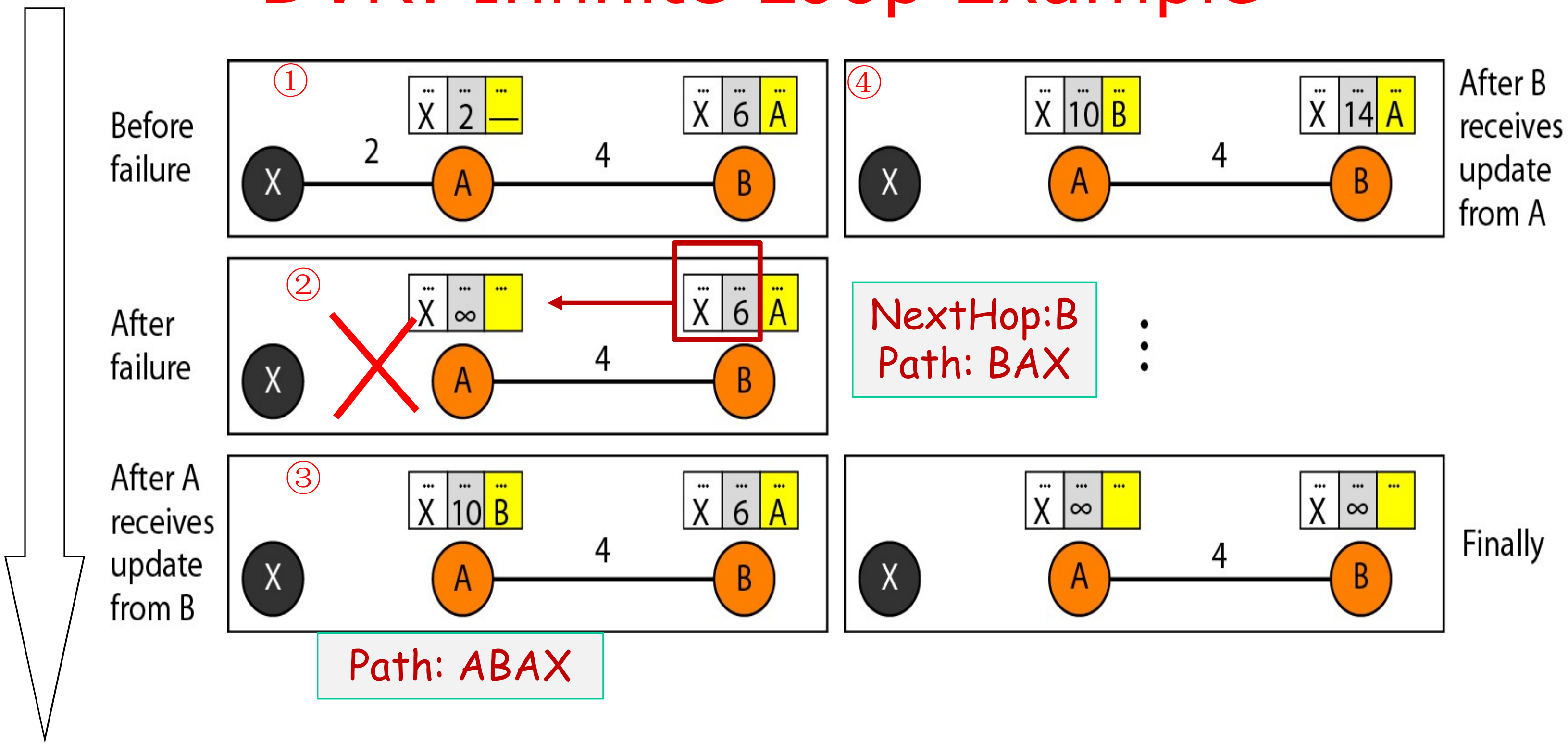
Good news: a new route is added

A	B	C	D	E	
	1	2	3	4	Initially
	3	2	3	4	After 1 exchange
	3	4	3	4	After 2 exchanges
	5	4	5	4	After 3 exchanges
	5	6	5	6	After 4 exchanges
	7	6	7	6	After 5 exchanges
	7	8	7	8	After 6 exchanges
		⋮			
	•	•	•	•	

(b)

Bad news （eg. a router is down） travels slowly

DVR: Infinite Loop Example



For a destination, router only knows distance, no idea of the path

(只知距离，不知路径)

DVR: Summary

- Used in
  - RIP (Routing Information Protocol)
  - Cisco EIGRP (Enhanced Interior Gateway Routing Protocol)
- Problem of DVR
  - When X tells Y that it has a path somewhere, Y has no way of knowing whether itself is on the path(whether there is a loop)
  - Rumor based routing
- Improvement: Path Vector Routing
  - BGP-4 (Border Gateway Protocol)
  - IDRP (Inter-domain Routing Protocol)

Review:

- Flooding（洪泛）：向除输入之外的所有端口转发，简单强壮、负载大
- Dijkstra：根据拓扑结构构建路由表
- 动态（自适应）的选路算法的前提：路由器要获知网络状况
  - 路由器之间需要交换信息
- 距离矢量选路（DVR）：邻居节点之间交换路由表
  - 只知距离、不知路径
  - 固有问题：无穷计数（无限环路）
  - 解决方法
    - 限制网络规模：RIP
    - 交换路径信息：BGP
    - 节点知道拓扑结构：LSR

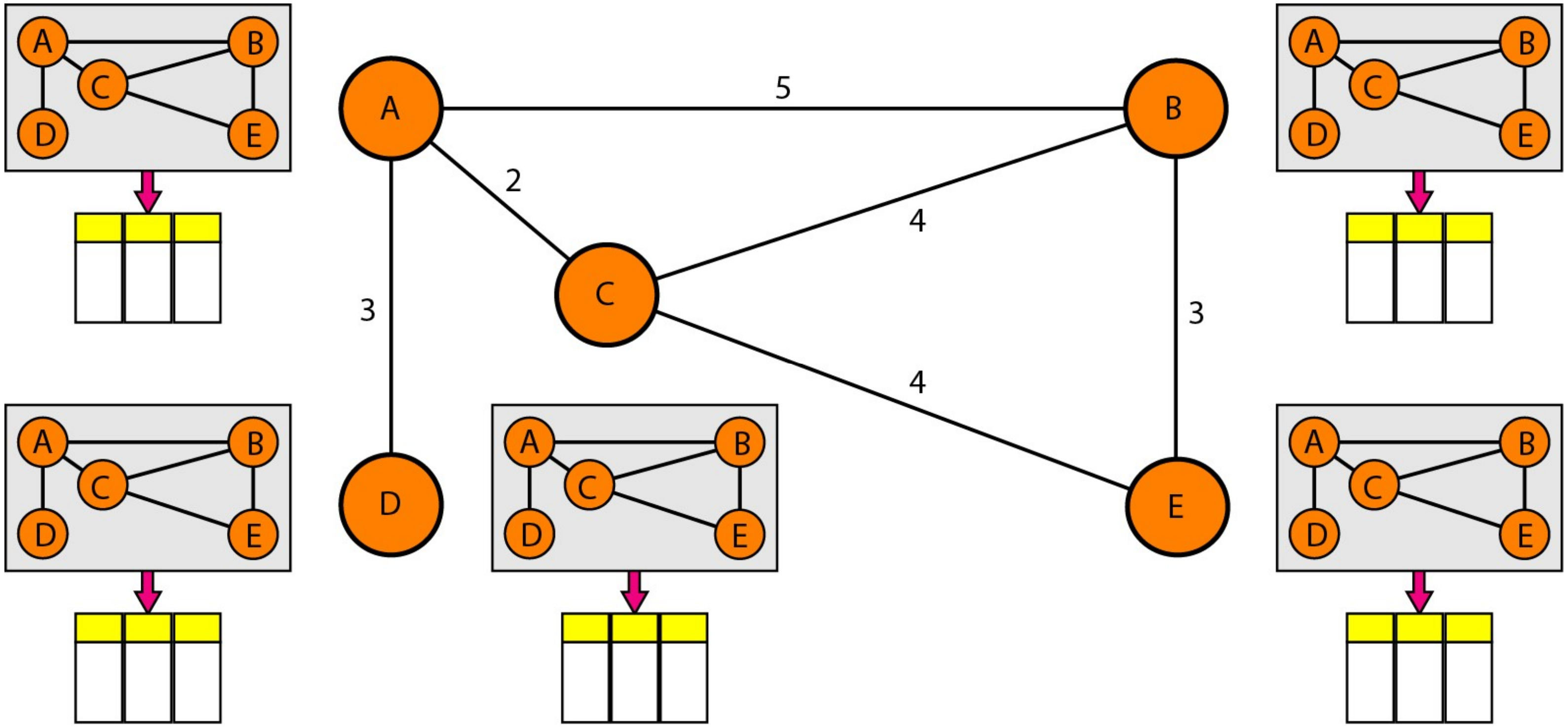
Outline

- Introduction to Network Layer
- *Routing Algorithms*
  - Concept of Routing and Routing Table
  - Optimality principle
  - Flooding
  - Dijkstra's "Shortest Path" Algorithm
  - Distance Vector Routing (DVR)
  - *Link State Routing (LSR)*
  - Hierarchical Routing
- Congestion Control and Internetworking
- The network layer in Internet

Link State Routing(链路状态选路)

- In LSR, each router shares its knowledge about the its neighbourhood with every other routers in the network.
- However, in LSR, the Link-State Packet (LSP) defines the best known network topology is sent to every routers after it is constructed locally.
- 3 keys to understand how LSR works:
  - Sharing knowledge about network topology.
  - Sharing with every other routers.
  - Sharing when there is a change.
- From the received LSPs and knowledge of entire topology, a router can then calculate the shortest path from itself to every other routers in the area.

LSR: What routers know?



The routers have knowledge of entire topology

Link State Routing: Operations

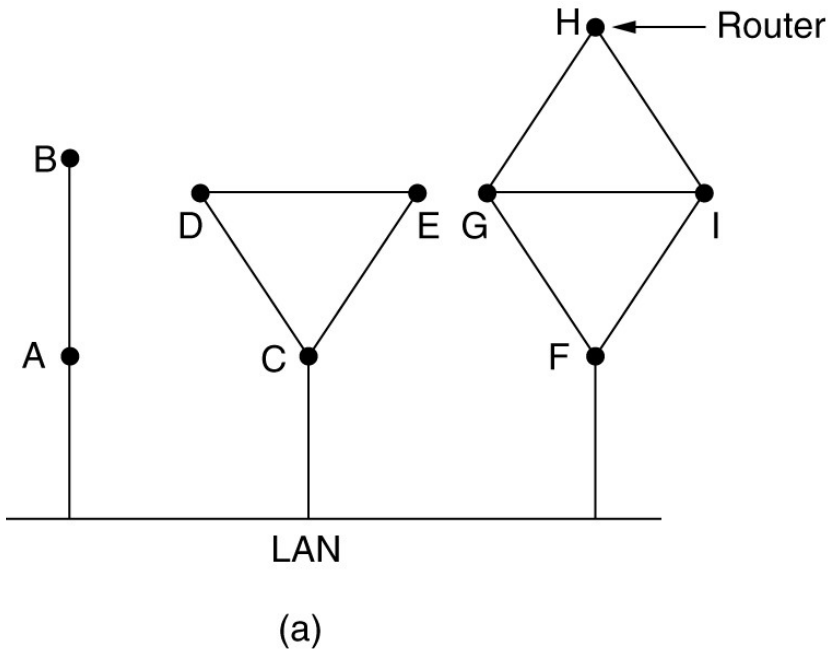
- Each router must do the following
  1. Discover its neighbors, learn their network address
  2. Measure the delay or cost to each of its neighbors
  3. Construct a packet (LSP: Link-state Packet, 链路状态包) telling all it has just learned
  4. Send LSP to all other routers (not only neighbors) reliably
  5. Compute the shortest path to every other router

Step1: Learning about the Neighbors

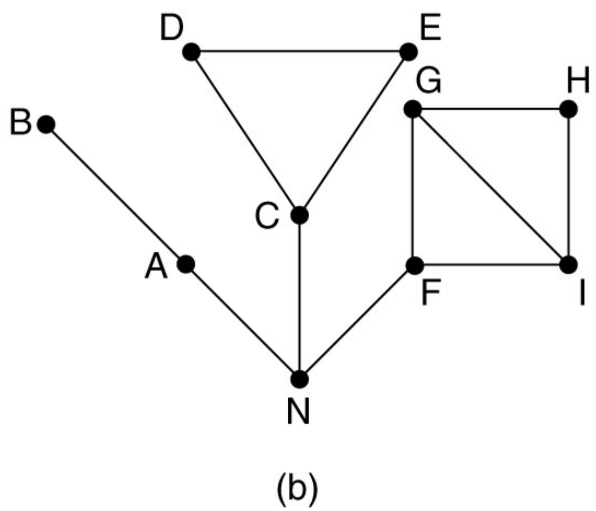
- HELLO packet
  - For a broadcast network, Broadcasting a special HELLO packet
  - For a point-to-point network, Unicast HELLO to neighbors
- Neighbors send back a reply telling who it is
  - Names (Route IDs) must be globally unique
- Simplify the topology
  - When two or more routers are connected by a LAN or other multi-access network

How to model the LAN

A LAN is considered as a Node(虚拟节点)



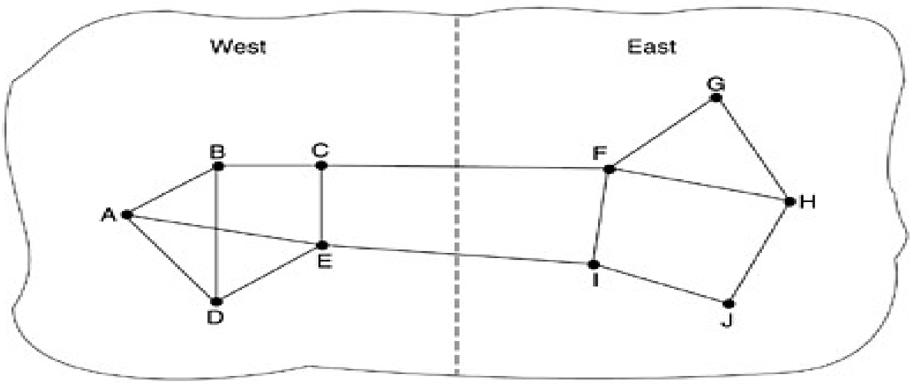
(a) Nine routers and a LAN.



(b) A graph model of (a)

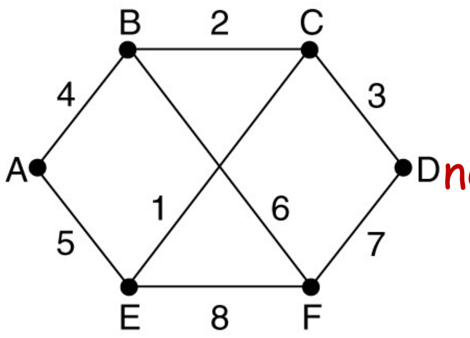
Step 2: Measuring Line Cost

- The delay to each of its neighbors
  - Measuring the round-trip time (ECHO Packet)
  - Bandwidth
  - An issue is whether to take the load into account when measuring the delay? (routing tables may oscillate, leading to erratic routing and many potential problems)



路由振荡，
导致不稳定

Step 3: Building Link State Packets



(a) A subnet.

Source
neighbor
cost

Link State Packets

A		B		C		D		E		F	
Seq.		Seq.		Seq.		Seq.		Seq.		Seq.	
Age		Age		Age		Age		Age		Age	
B	4	A	4	B	2	C	3	A	5	B	6
E	5	C	2	D	3	F	7	C	1	D	7
		F	6	E	1			F	8	E	8

(b) The link state packets for this subnet

When to build LSP

- Build LSP periodically
- Build LSP when some significant event occurs
  - such as a line or neighbor going down or coming back up again or changing its properties appreciably

LSP: Why using Sequence number & Age?

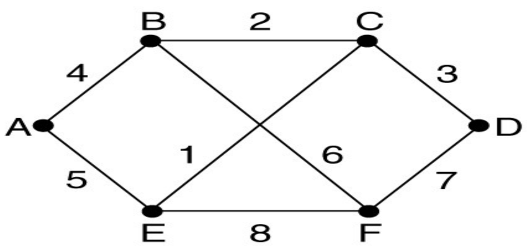
- Sequence number: used to check whether the received LSP is new
  - If LSP is new, forward with flooding
  - If LSP is a duplicate, it is discarded
  - If lower than the highest one seen so far ever arrives, it is rejected as being obsolete
  - If a router is down, and then restart, ...
- Age: used to delete looped LSP
  - If a router is down
  - If a router changed its routerID
  - If a router need to delete an LSP

Step 4: Distributing the Link State Packets

- Responsibility
  - Distributing the LSP reliably

- The basic distribution algorithm
  - Use flooding to distribute LSPs
  - Guard against errors on the router-router lines, all LSP are acknowledged

LSP: Example



(a)

Link State Packets

A	
Seq.	
Age	
B	4
E	5

B	
Seq.	
Age	
A	4
C	2
F	6

C	
Seq.	
Age	
B	2
D	3
E	1

D	
Seq.	
Age	
C	3
F	7

E	
Seq.	
Age	
A	5
C	1
F	8

F	
Seq.	
Age	
B	6
D	7
E	8

(b)

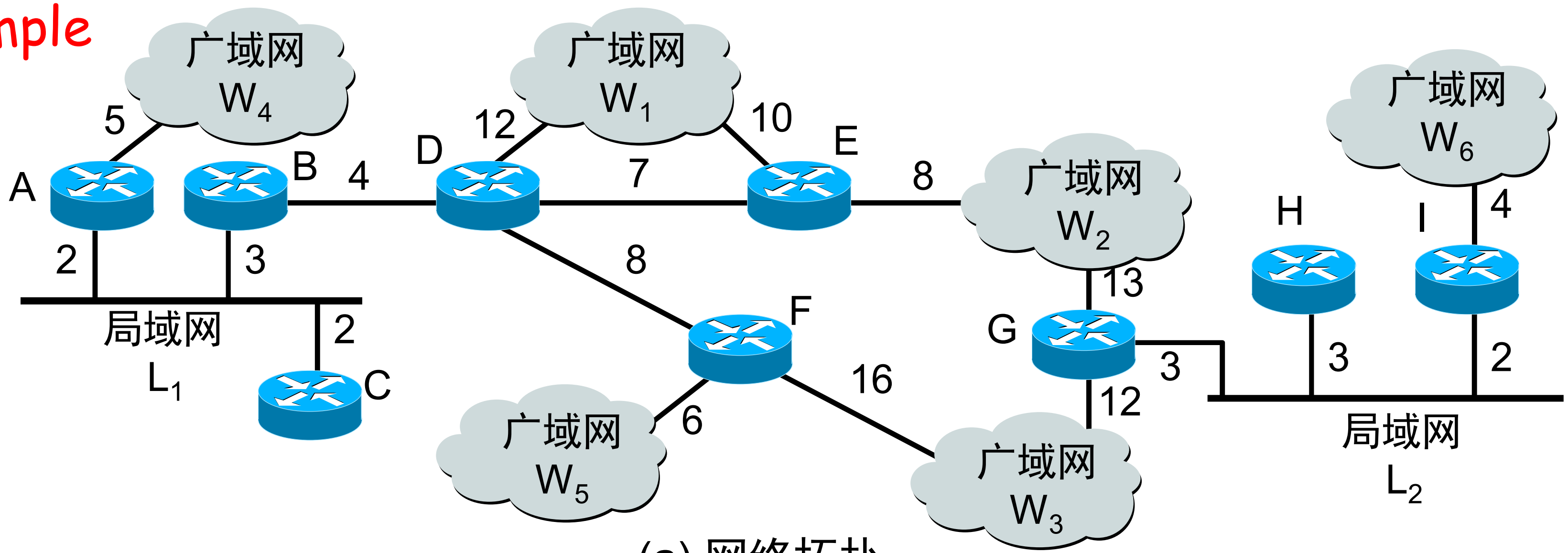
Record of LSPs received by B

Source	Seq.	Age	Send flags A	Send flags C	Send flags F	ACK flags A	ACK flags C	ACK flags F	Data
A	21	60	0	1	1	1	0	0	
F	21	60	1	1	0	0	0	1	
E	21	59	0	1	0	1	0	1	E 发出的LSP
C	20	60	1	0	1	0	1	0	
D	21	59	1	0	0	0	1	1	

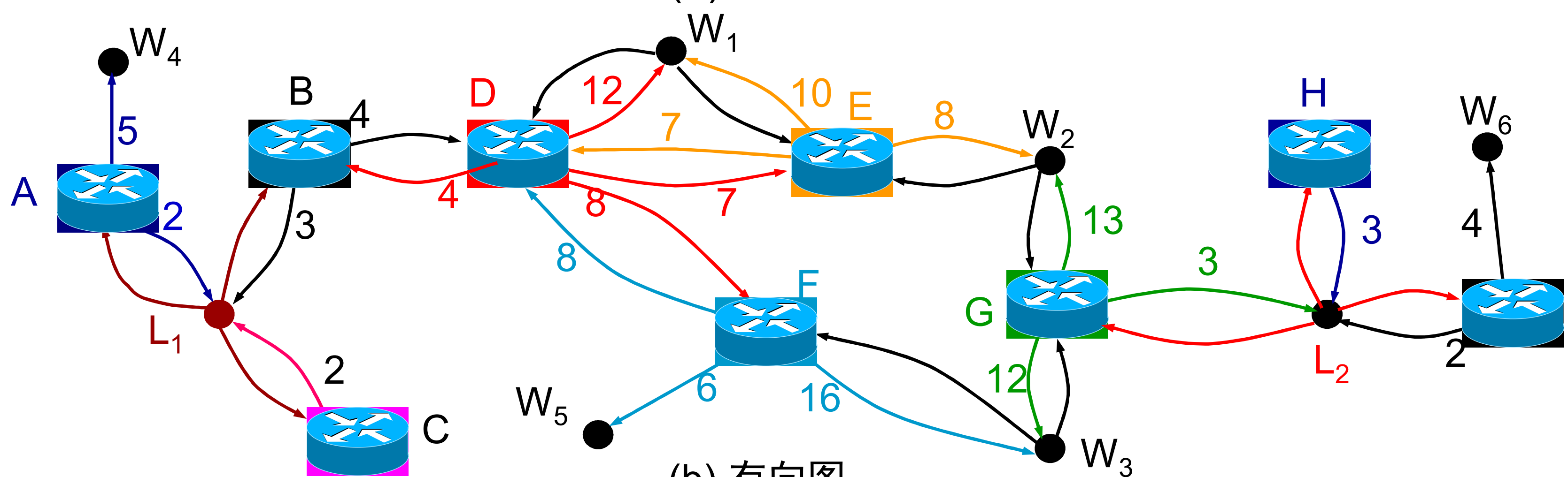
Step 5: Computing the New Route

- Once a router has accumulated a full set of LSPs, it can construct the entire graph
  - Every link is represented twice, once for each direction. The two values can be averaged or used separately.
- Run Dijkstra's algorithm locally to construct the shortest path to all possible destinations
- The results of this algorithm can be installed in the routing tables

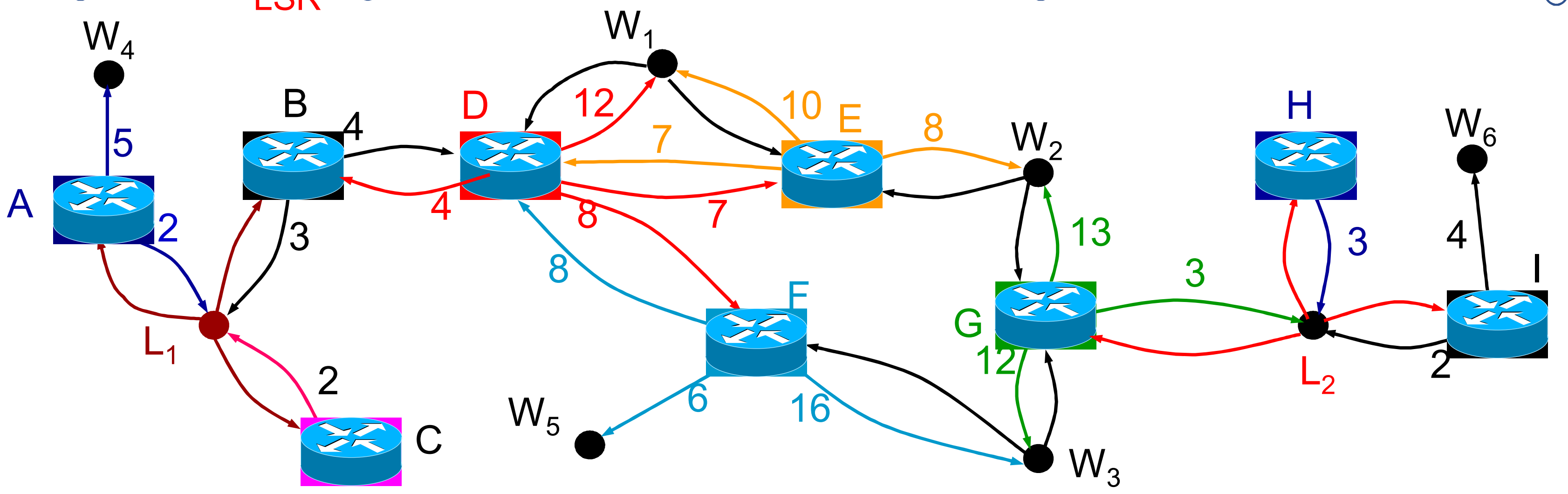
LSR: Example



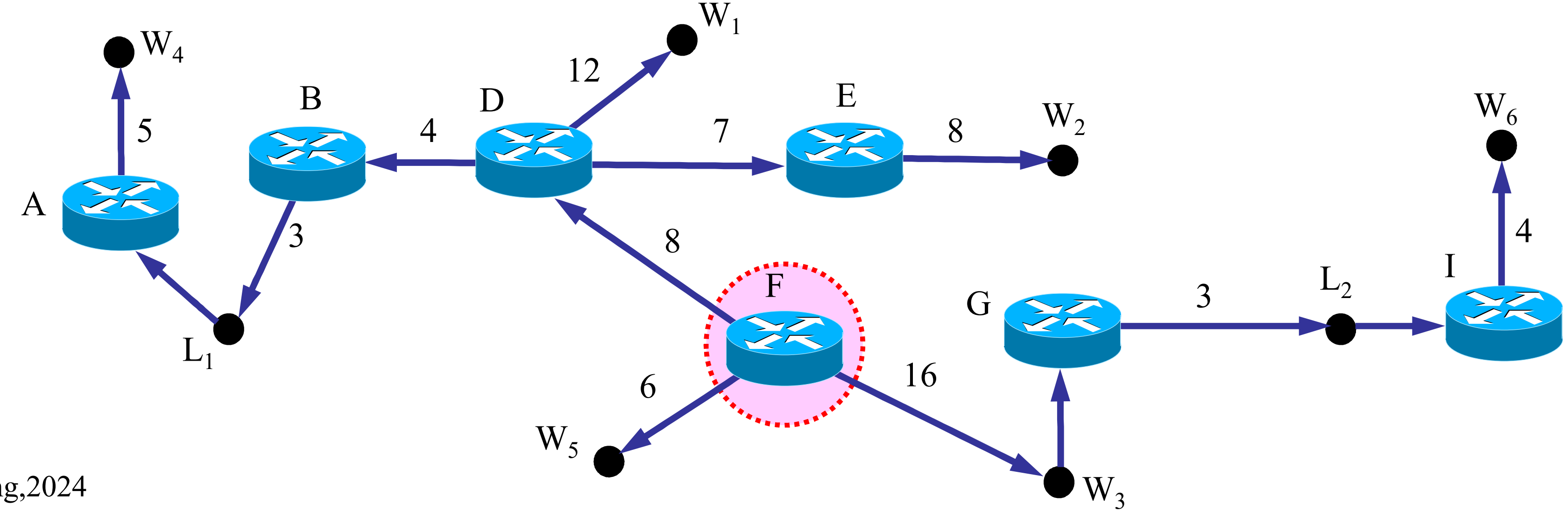
(a) 网络拓扑



(b) 有向图



以路由器F为根的最短路径树



F的路由表

Dest	NextHop
L1	D
L2	G
W1	D
W2	D
W3	-- (direct)
W4	D
W5	-- (direct)
W6	G

Adoptions of Link State Routing

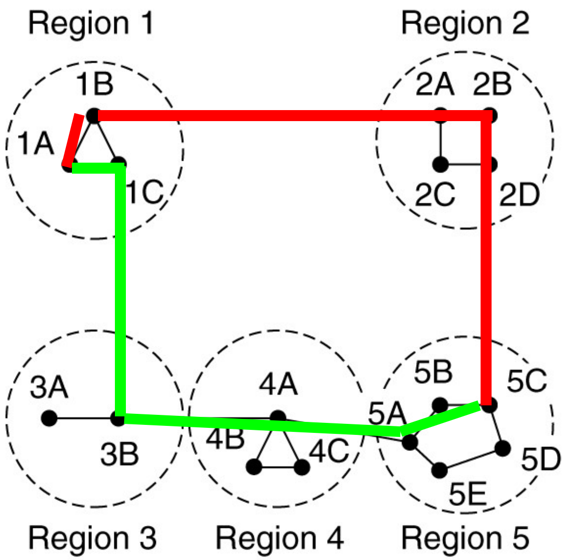
- IS-IS (Intermediate System-Intermediate System):
  - support multiple network layer protocols at the same time
  - DECnet, ISO CLNP, IP, AppleTalk, Novell NLSP (IPX)

- OSPF (Open Shortest Path First)
  - designed several years after IS-IS
  - Only for IP

Outline

- Introduction to Network Layer
- ***Routing Algorithms***
  - Concept of Routing and Routing Table
  - Optimality principle
  - Flooding
  - Dijkstra's "Shortest Path" Algorithm
  - Distance Vector Routing (DVR)
  - Link State Routing (LSR)
  - *Hierarchical Routing*
- Congestion Control and Internetworking
- The network layer in Internet

Hierarchical Routing(层次选路/分级选路) (1)



1A→5C

(a)

Full table for 1A

Dest.	Line	Hops
1A	–	–
1B	1B	1
1C	1C	1
2A	1B	2
2B	1B	3
2C	1B	3
2D	1B	4
3A	1C	3
3B	1C	2
4A	1C	3
4B	1C	4
4C	1C	4
5A	1C	4
5B	1C	5
5C	1B	5
5D	1C	6
5E	1C	5

(b)

Hierarchical table for 1A

Dest.	Line	Hops
1A	–	–
1B	1B	1
1C	1C	1
2	1B	2
3	1C	2
4	1C	3
5	1C	4

(c)

May take sub-optimal route (次优路由)

Hierarchical Routing: Example

Consider a subnet with 720 routers.

- If there is no hierarchy
  - 720 routing table entries
- If the subnet is partitioned into 24 regions of 30 routers each (Total:53)
  - 30 local entries
  - 23 remote entries
- If a three-level hierarchy is chosen, with 8 clusters, each containing 9 regions of 10 routers (Total: 25)
  - 10 entries for local routers
  - 8 entries for routing to other regions within its own cluster
  - 7 entries for distant clusters

Summary of Routing Algorithms

- Distance Vector Routing vs Link State Routing

	DVR	LSR
What information to share	routing table(destination and distance)	network topology (LSP)
Whom to share with	neighbors	All routers in the network
When to share	periodically	When there is change
What does router knows	distance, not path	Entire network topology
Count-to-infinity problem?	Yes	No

- Hierarchical Routing(分级选路): 简化路由表
  - 域内：选路到路由器；域间：选路到域