Chapter5 Network Layer(3)

王昊翔 hxwang@scut.edu.cn

School of Computer Science & Engineering ,SCUT

Communication & Computer Network key-Lab of GD

Contents of the lecture

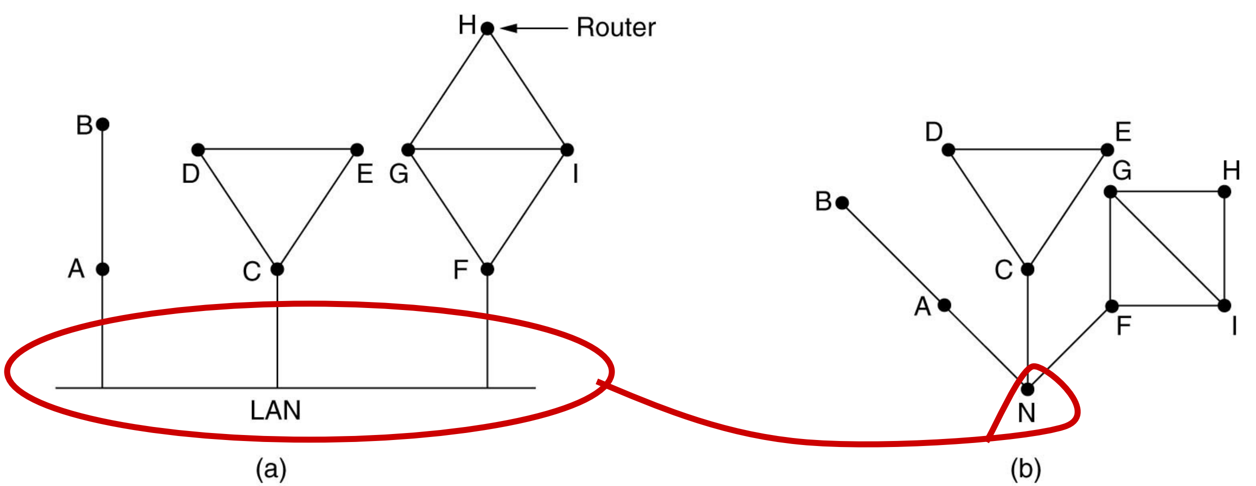
- Review of last lecture
  - DV
  - Problem
- Link state routing algorithm
  - An example： OSPF
- BGP

Link State Routing

- Distance vector routing was used in the ARPANET until 1979, when it was replaced by link state routing.
- Variants of link state routing are now widely used.
- The idea behind link state routing consists of five parts:
  - Discover its neighbors and learn their network addresses.
  - Measure the delay or cost to each of its neighbors.
  - Construct a packet telling all it has just learned.
  - Send this packet to all other routers.
  - Compute the shortest path to every other router.

Learning about the Neighbors

- When a router is booted, it sends a special HELLO packet on each point-to-point line.
- The router on the other end is expected to send back a reply telling who it is (using a globally unique name).
- When two or more routers are connected by a LAN, the LAN can be modeled as a node.

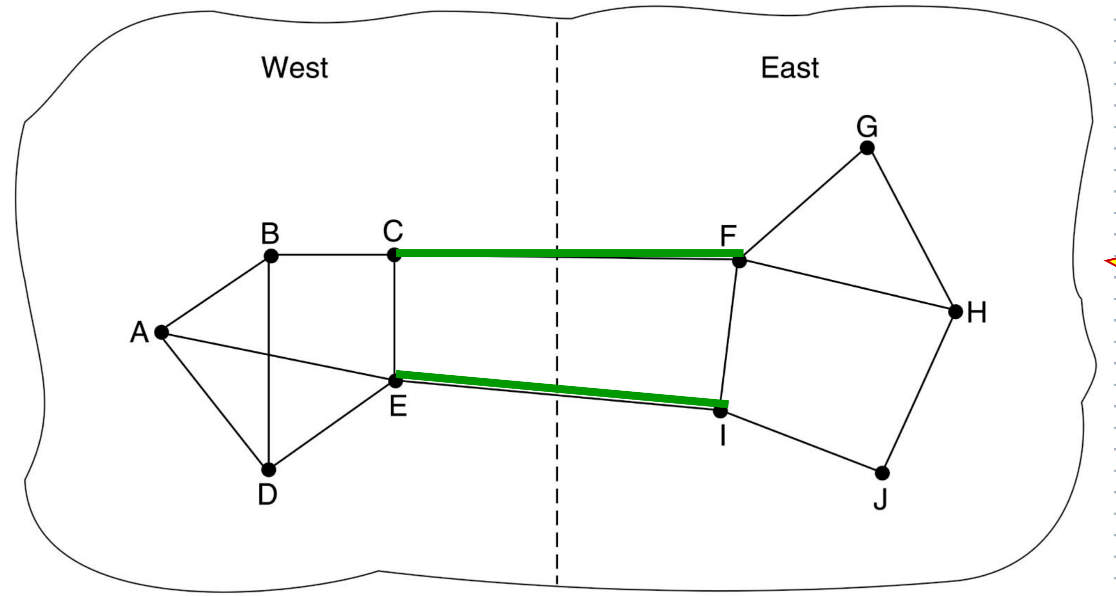


Measuring Line Cost

- To determine the cost for a line, a router sends a special ECHO packet ,and that the other side is required to send back immediately.
- By measuring the round-trip time, the sending router can get a reasonable estimate of the delay.
  - For even better results, the test can be conducted several times, and the average used.
- To factor the load in, the round-trip timer must be started when the ECHO packet is queued.
- To ignore the load, the timer should be started when the ECHO packet reaches the front of the queue.

Measuring Line Cost (cont'd)

- **Should the load be taken into account when measuring the delay ?**
  - **Arguments can be made both ways.**



Should be distribute load

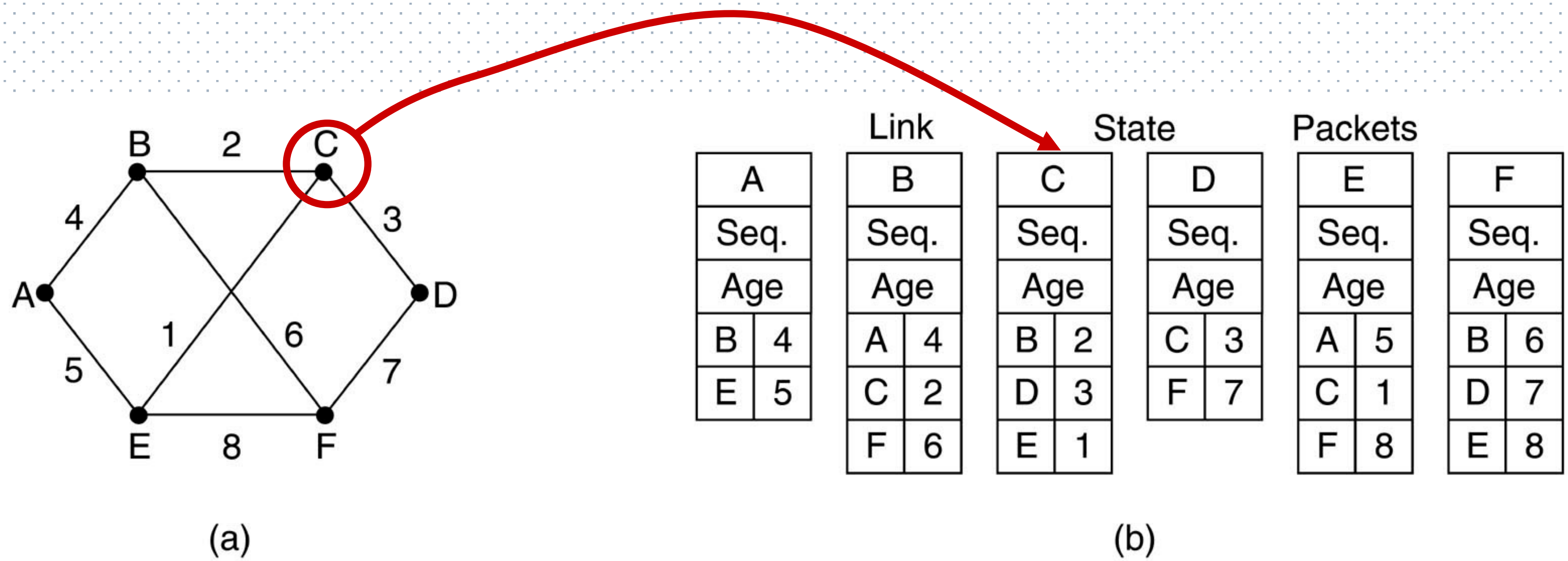
Building Link State Packets

- A Link State Packet is constructed to send to other routers. Information contained in the packet is:
  - ID of the sender
  - sequence number
  - age
  - list of neighbors
  - delay to each neighbor

Link state information

- when to build?
  - State packets may be built periodically, or when some significant event occurs, such as a line or neighbor going down or coming back up again.

Building Link State Packets (cont'd)



(a) A subnet. (b) The link state packets for this subnet.

Distributing The Link State Packets

- The basic algorithm:
  - Each state packet contains a sequence number that is incremented for each new packet sent.
  - Routers keep track of all the (source router, sequence) pairs they see.
  - When a new link state packet comes in, it is checked against the list of packets already seen.
    - If it is new, it is forwarded on all lines except the one it arrived on (i.e., flooding).
    - If it is a duplicate, it is discarded.
    - If a packet with a sequence number lower than the highest one seen so far ever arrives, it is rejected as being obsolete.

Distributing The Link State Packets (cont'd)

- **Problems with the basic algorithm:**
  - **The sequence numbers may wrap around, causing confusion.**
    - **Solution: using a 32-bit sequence number. With one packet per second, it would take 137 years to wrap around.**
  - **If a router ever crashes, it will lose track of its own sequence number. If it starts again at the sequence number 0, new packets will be rejected as obsolete/duplicate by other routers.**
  - **If a sequence number is ever corrupted and 65,540 is received instead of 4 (a 1-bit error), packets 5 -- 65540 will be rejected as obsolete.**

00000000000000100
10000000000000100

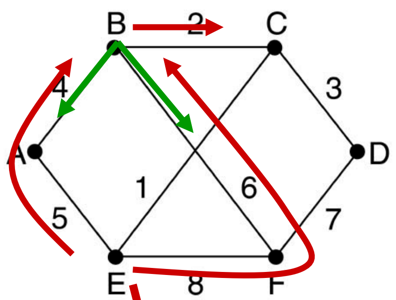
Distributing The Link State Packets (cont'd)

- The solution to router crashes and sequence number corruption is to associate an age (e.g., 60) with each state packet from any router and decrement the age once per second.
- When the age hits zero, the information from that router is discarded.
- Normally a new packet comes in every 10 seconds, so router information only times out when a router is down (or 6 consecutive packets have been lost, an unlikely event).

Distributing The Link State Packets (cont'd)

- Some refinements to the basic algorithm make it more robust.
  - When a state packet comes in to a router for flooding, it is put in a holding area (保留区)to wait a short while first.
  - If another state packet from the same source comes in before it is transferred, their sequence numbers are compared.
    - If they are equal, the duplicate is discared.
    - If they are different, the older one is thrown out.
  - To guard against errors on the lines, all state packets are acknowledged.
  - When a line goes idle(空闲), the holding area is scanned in round robin to select a packet or acknowledgement to send.

Distributing The Link State Packets (cont'd)



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

C's LSP arriv. From F

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	
C	20	60	1	0	1	0	1	0	
D	21	59	1	0	0	0	1	1	

1 0 0 0 1 1

B's holding area

Computing the New Routes

- A full set of link state packets allows a router to construct a graph of the entire subnet.
- We can now use Dijkstra's algorithm to figure out the shortest paths between routers.
- We can install this information in the routers to direct the packets. (set up routing-table)

Characteristics of L-S routing algorithm

- **Advantages**
  - **Consistency of every router is good**
  - **Convergence is good**
  - **Fit for big network**

- **Disadvantages**
  - **Each router requires bigger storage-space**
  - **Computing workload is great**

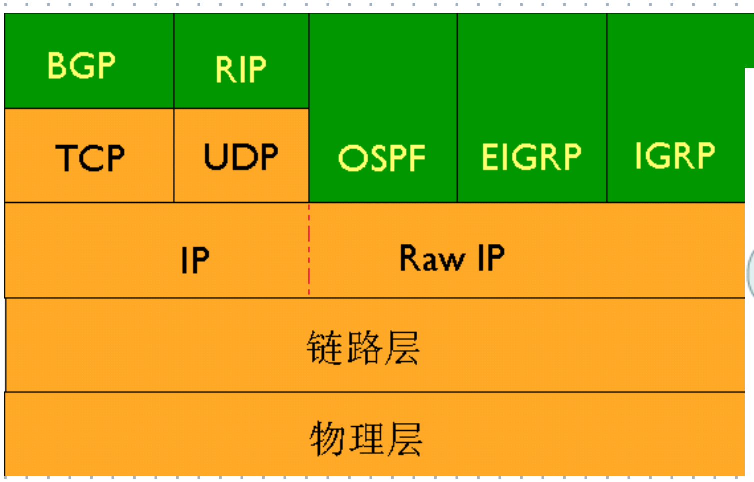
Example of L-S routing protocol—OSPF

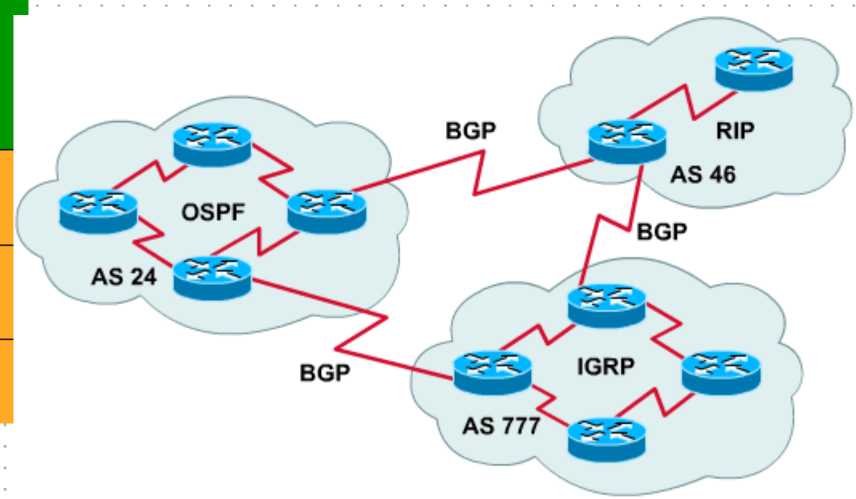
- Group talk

- Open shortest path first
- Use graph to replace real network
  - Every router is a node
  - Measure cost（metric）
  - May have a few graphs
- Computing shortest path

BGP（border gateway protocol）（边界网关协议）

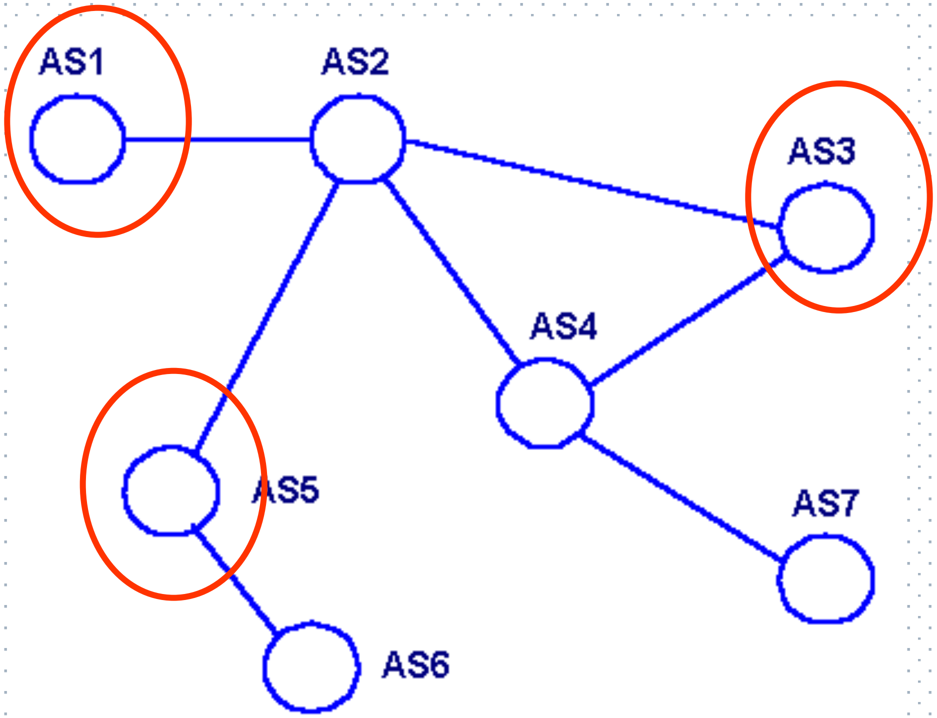
- different protocol - BGP (Border Gateway Protocol) is needed between ASes because the goals of an interior gateway protocol and an exterior gateway protocol are not the same.
- The definition of BGP is in RFCs 1771 to 1774.





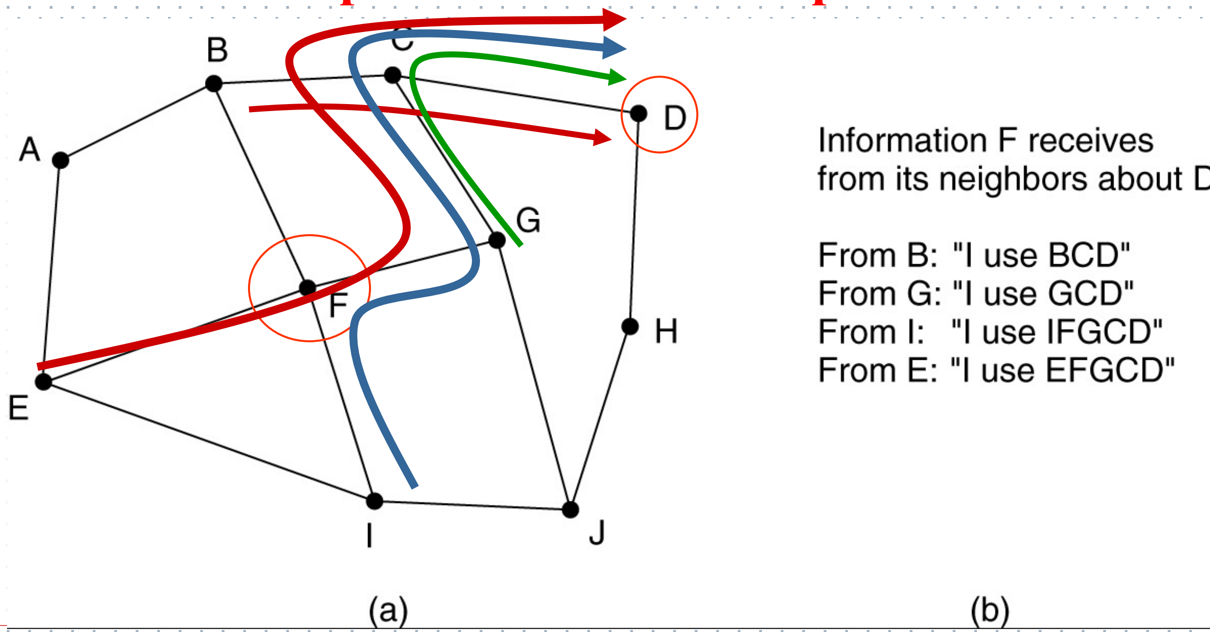
BGP Principle（1/2）

- The typical policies of exterior gateway protocol routers involve political, security, or economic considerations.
- Given BGP's special interest in transit traffic, networks are grouped into one of three categories.
  - stub networks
  - multiconnected networks
  - transit networks


AS1
AS2
AS3
AS4
AS5
AS6
AS7

BGP原理（2/2）P459~461

- Pairs of BGP routers communicate with each other by establishing TCP connections.
- BGP is fundamentally a distance vector protocol, but quite different from most others such as RIP.
  - BGP router **keeps track of the exact path**.



Summary

- Link state algorithm
  - Five steps
  - Problems and their resolution
- OSPF
  - Five message types
  - DR selection
  - OSPF operation process(status)
- BGP

Thank you all !