Communication Networks:
Technology & Protocols

Stavros Tripakis (stavros@eecs)

Lecture 12

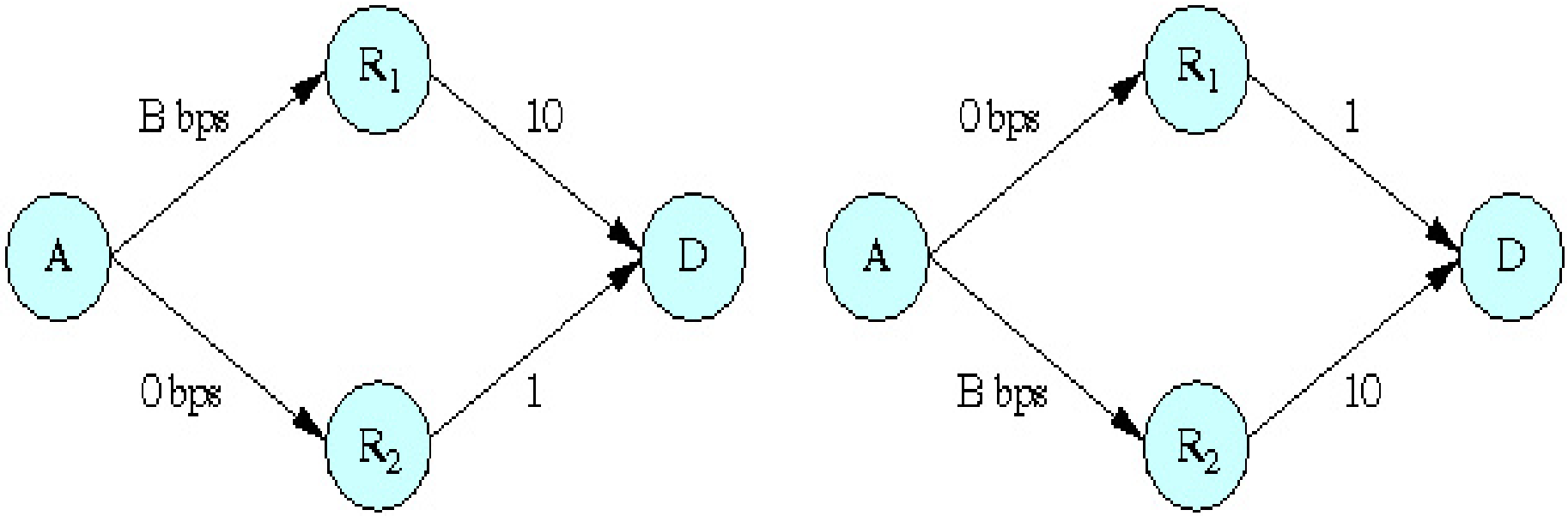
September 19

Routing metrics

- How are link weights determined ?
  - Simple: weight 1 to all links (cost of path = hop count).
  - Link latency (queuing + propagation delays).
  - Link capacity (bit-rate).
  - Link reliability (packet-loss probability).
- Routing protocols might use multiple metrics and compute multiple routing tables for different traffic requirements (OSPF does this).
- Problem: link state changes dynamically.

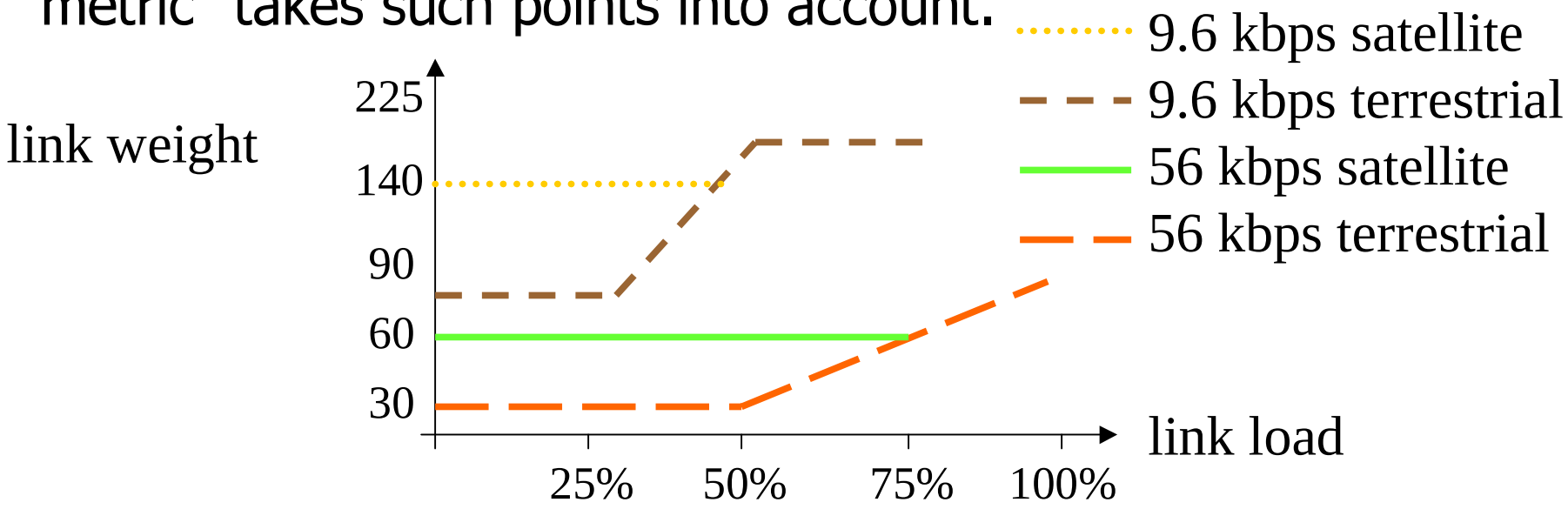
Routing metrics: example of routing oscillations

- Routing affects link load.
- Link load affects link weight (latency, available capacity, etc).



Routing metrics

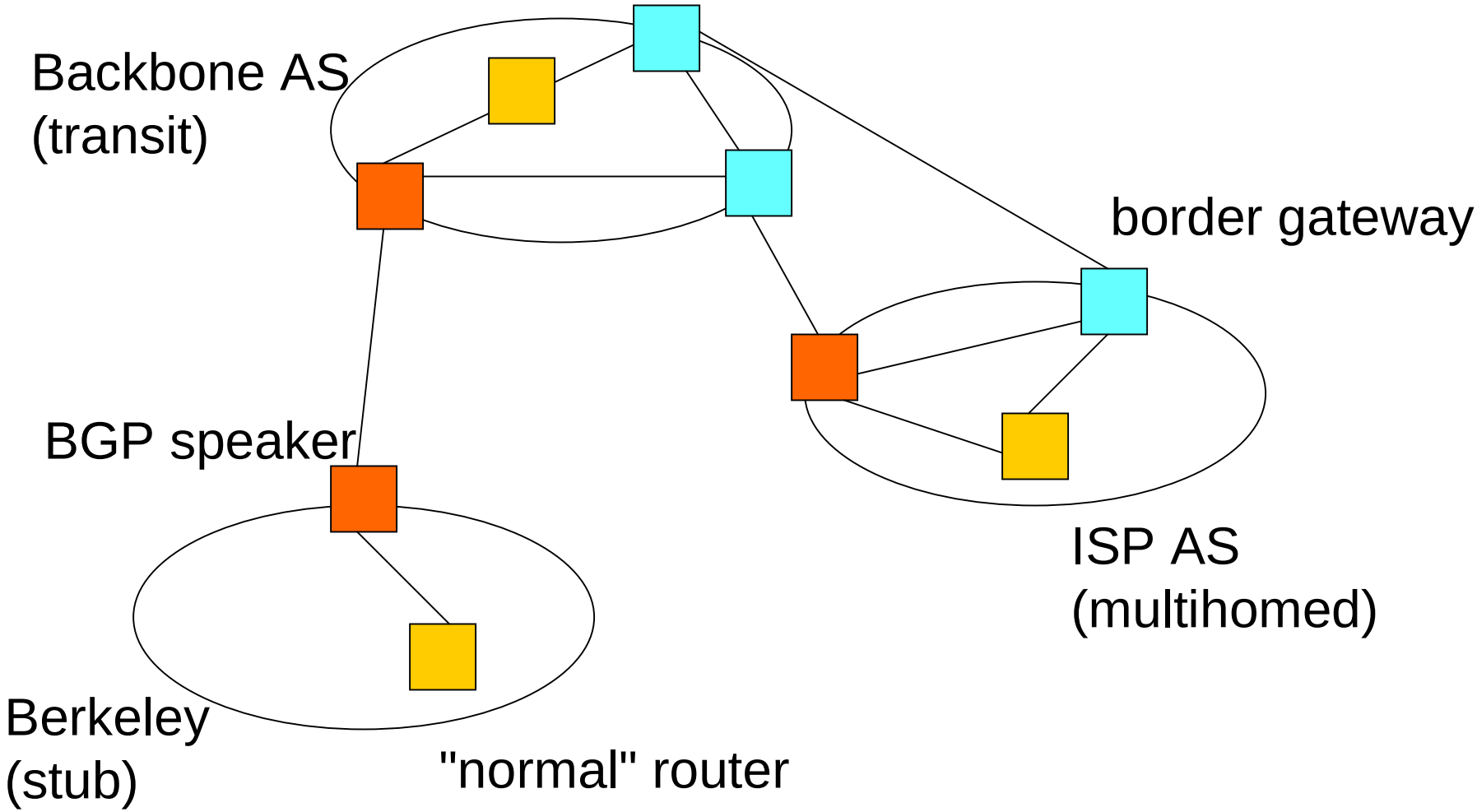
- To avoid such oscillations:
  - "Normalize" different routing metrics.
  - Take into account link type (e.g., satellite, terrestrial).
  - Smoothen the variation of metric in time.
- (After many experiments) "Revised ARPANET routing metric" takes such points into account.



Interdomain routing

- At the IP (network) level.
- Network: set of AS connected by **border gateways** (routers connecting one AS to another).
- AS classification:
  - *stub* AS: single connection to another AS, local traffic;
  - *multihomed* AS: multiple connections, local traffic;
  - *transit* AS: multiple connections, transit traffic.
- One border gateway from each AS is the BGP **speaker**.

Interdomain routing



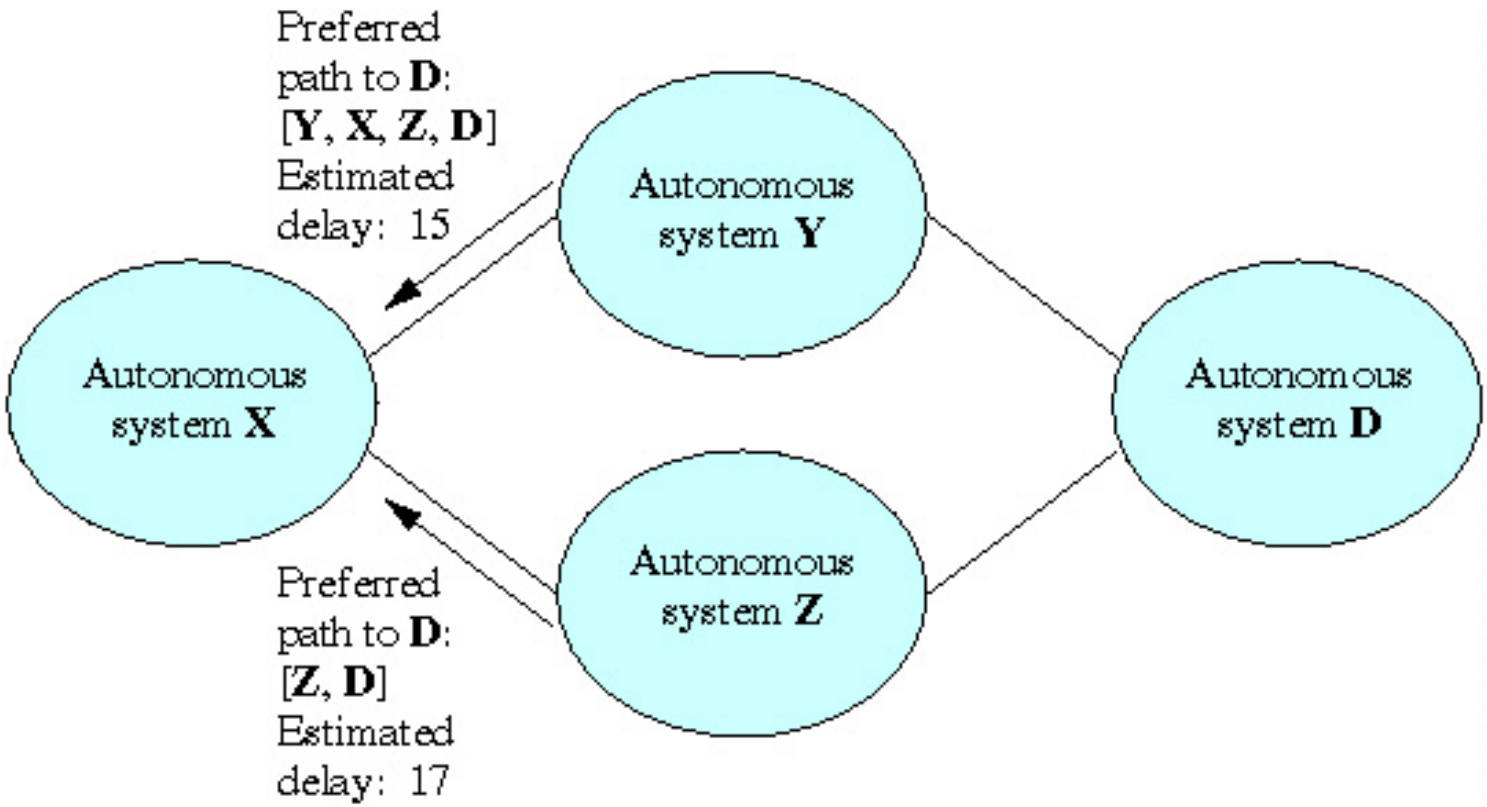
Interdomain routing: BGP

- BGP (border gateway protocol) is executed among BGP speakers.
- BGP speakers then distribute routing information to other border gateways inside their AS.
- Why not execute OSPF instead of BGP ?
  - Too many ASs (50,000): flooding becomes too expensive.
  - Administrative reasons: security, economics (cf. stub vs. transit ASs).
  - Metrics not always consistent among ASs.

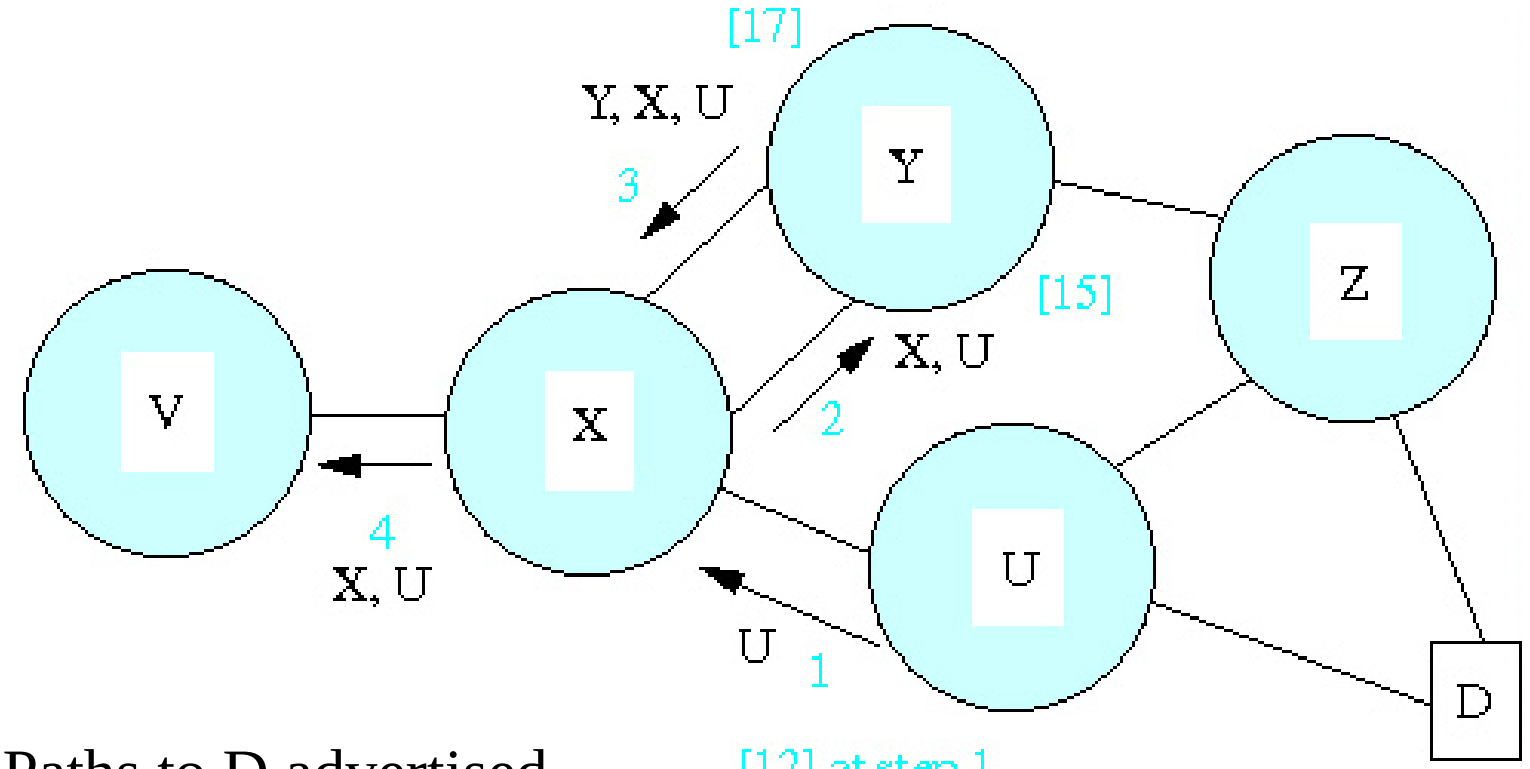
Interdomain routing: BGP

- BGP speakers advertise **preferred** paths (routes).
- BGP is distributed: each speaker has its own view of the network (might be inconsistent w.r.t. the view of other speakers).
- To avoid loops, advertise **complete** paths, e.g., AS X says: "My preferred path to AS Z is (X, Y, W, Z)".
- When Y hears this, it knows it shouldn't go through X to get to Z.

BGP: example


Preferred
path to **D**:
[Y, X, Z, D]
Estimated
delay: 15
Autonomous
system **Y**
Autonomous
system **X**
Autonomous
system **D**
Autonomous
system **Z**
Preferred
path to **D**:
[Z, D]
Estimated
delay: 17

BGP: another example

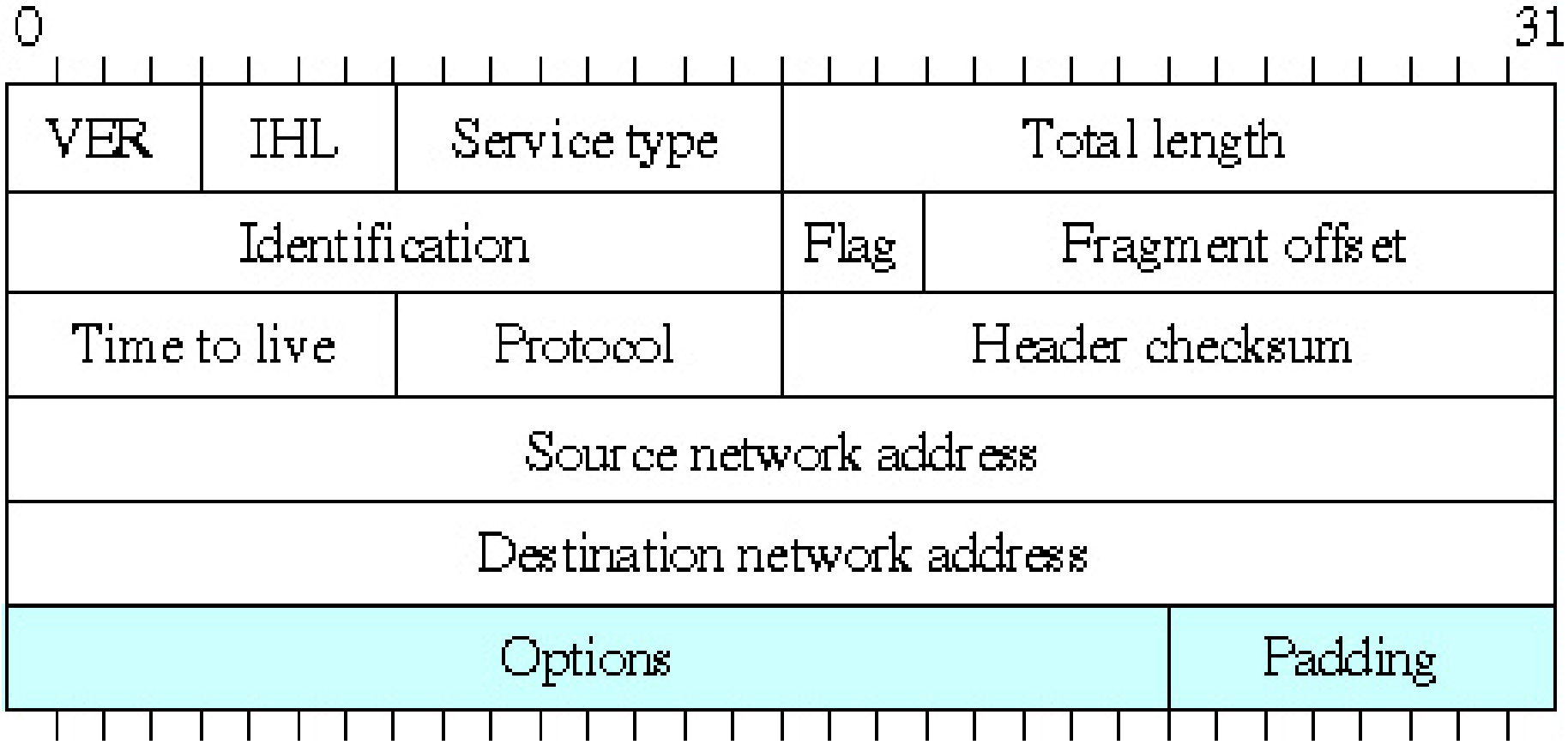


Note: Paths to D advertised
from Y and U to X
at time 3 are inconsistent.

IP (Internet Protocol)

- Network-layer protocol: unreliable **end-to-end** delivery of packets (**datagrams**) of up to 64 kbytes.
- End-to-end: source and destination might not be directly connected, routing involved.
- Unreliable: packets might be lost (corrupted at the physical level or dropped because of full buffers) or not be delivered successfully (destination unreachable, loops). **ICMP** informs the source in the latter case.

IP header format



IP fragmentation/reassembly

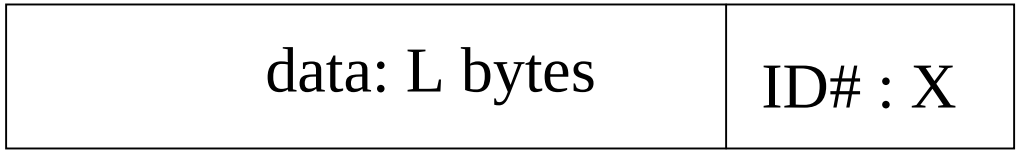
- Different link layers have different **Maximum Transfer Units** (MTUs): maximum size of packet they can transmit.
- When *Size(datagram) > MTU* the datagram needs to be split in many pieces: **fragmentation**.
- The pieces are joined together to form the original datagram at the destination: **reassembly**.
- A datagram may be fragmented multiple times along its route to the destination.

IP fragmentation/reassembly

- ID# used to identify the datagram: all fragments of the same datagram have same ID#.
- Offset used to mark the "starting position" of the fragment from the beginning of the datagram (in number of bytes).

IP fragmentation/reassembly

- Example:



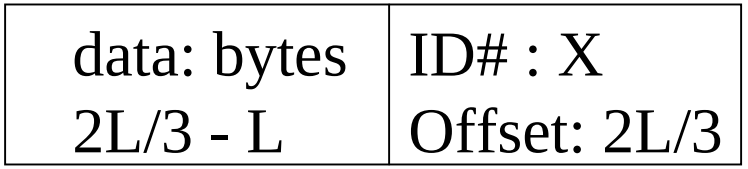
original datagram

fragment 1

MTU ≈ L/3

fragment 2

fragment 3



Modern internetworking

- Mobility:
  - Plug-and-play: allowing hosts to dynamically/temporarily connect to the Internet, in a particular domain. Dynamic Host Configuration Protocol (DHCP).
  - Allowing hosts to move between multiple domains. Mobile IP.
- Multicasting: sending packets to multiple hosts.
  - Simple (inefficient) solutions: flooding, multiple transmissions.
  - Spanning tree, Multicast OSPF, PIM, etc.