Computer Networks and Internets

《计算机网络与因特网》课件

林坤辉

PART IV Internetworking

Chapter 22（2）

IP Encapsulation, Fragmentation, and Reassembly

IP封装、分段与重组

22.2 Datagram Transmission and Frames

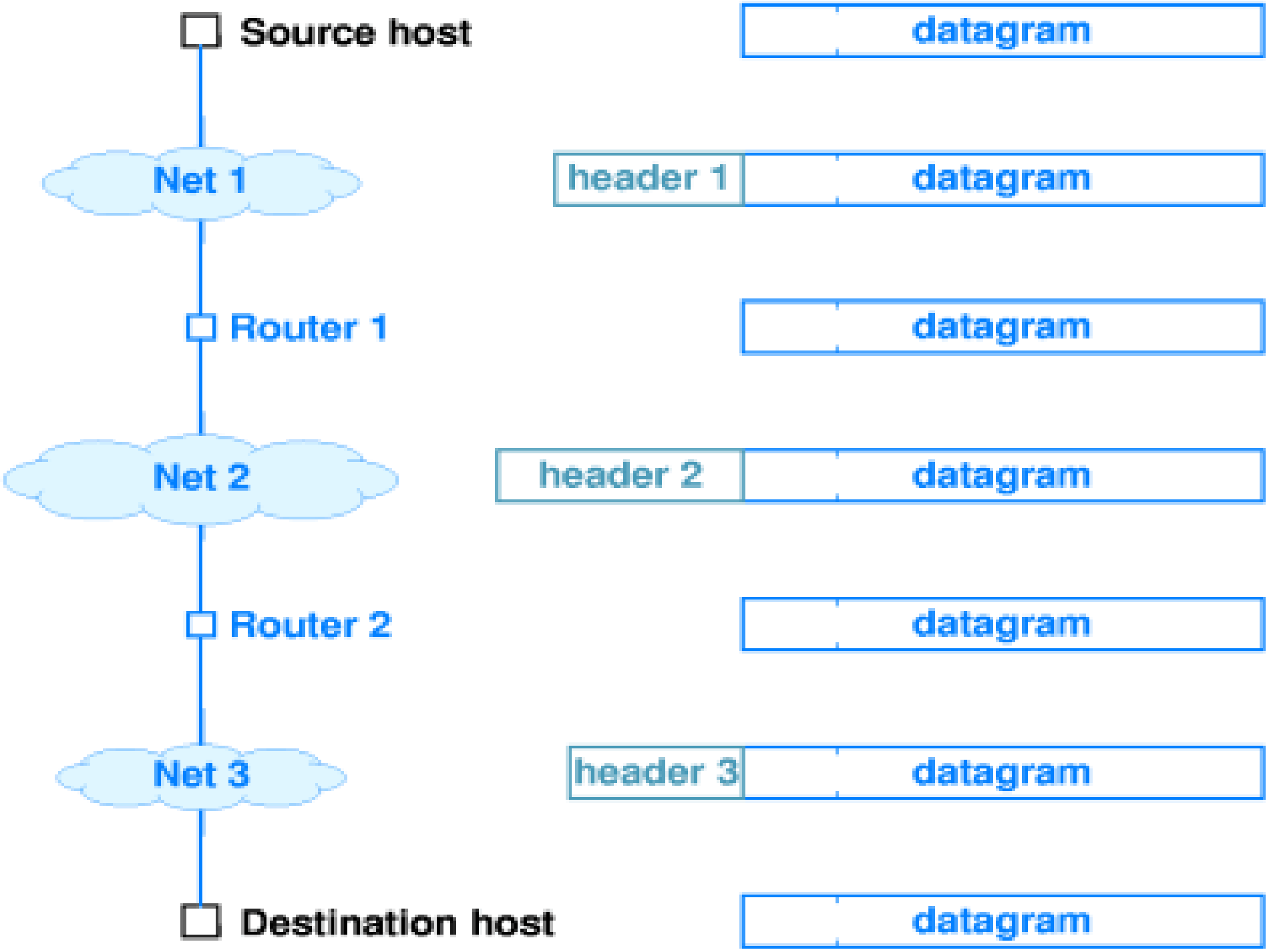
- IP software first selects the next hop,
- Then transmits the datagram across a physical network.
- Network hardware does not understand datagram format or Internet addressing.
- Each hardware technology defines a frame format and a physical addressing scheme.

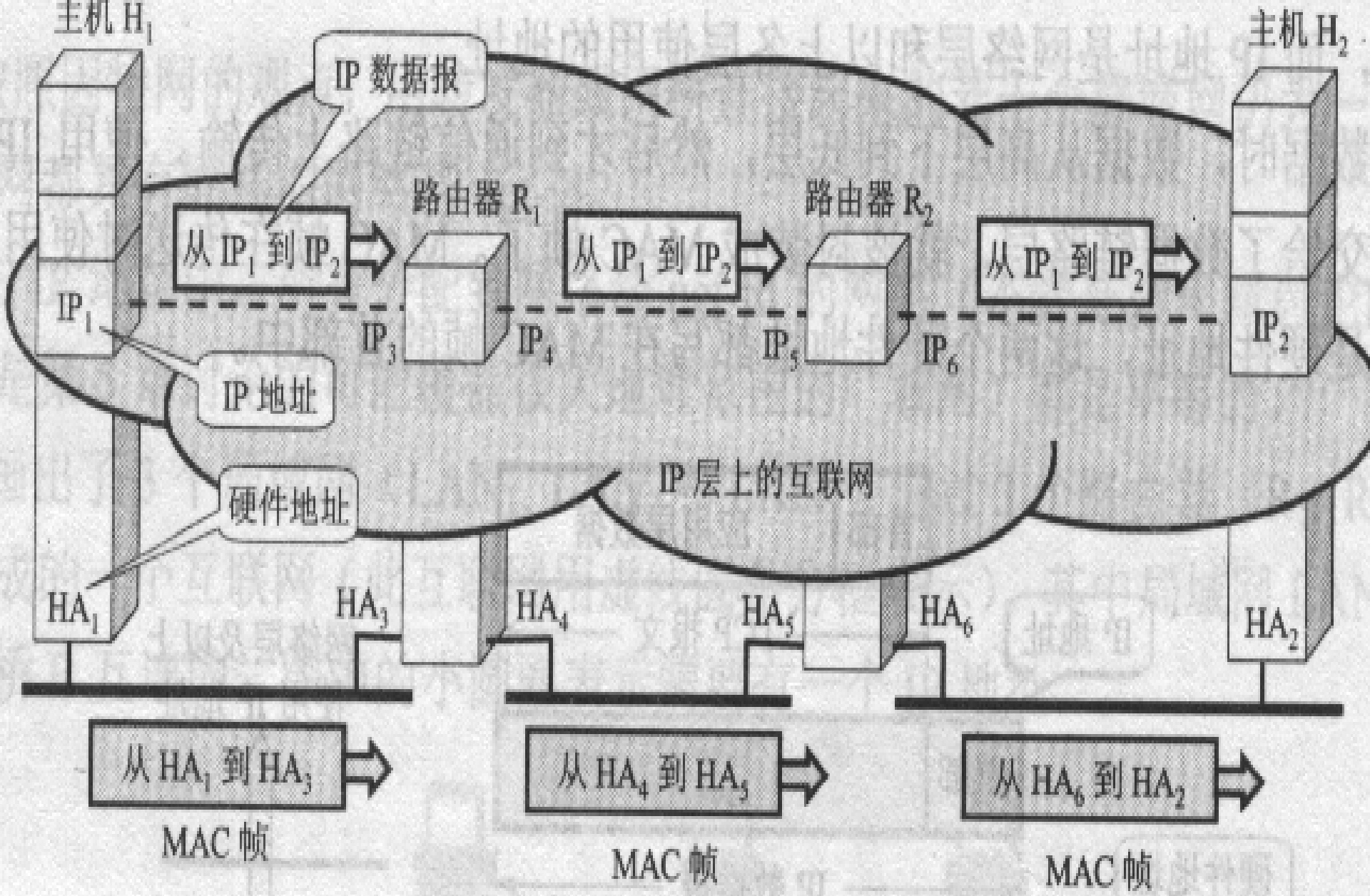
22.3 Encapsulation封装

- encapsulation.
- The hardware does not examine or change the contents of the frame data area.
- The sender and receiver must agree on the value used in the frame type field.
- The sender must supply the physical address of the next computer to which the datagram should be sent.

22.4 Transmission Across an Internet

- When a frame reaches the next hop,the receiving software extracts the IP datagram and discards the frame header.
- If the datagram must be forwarded across another network , a new frame is created.
- Frame headers do not accumulate during a trip through the internet.
- Hosts and routers store a datagram in memory with no additional header.

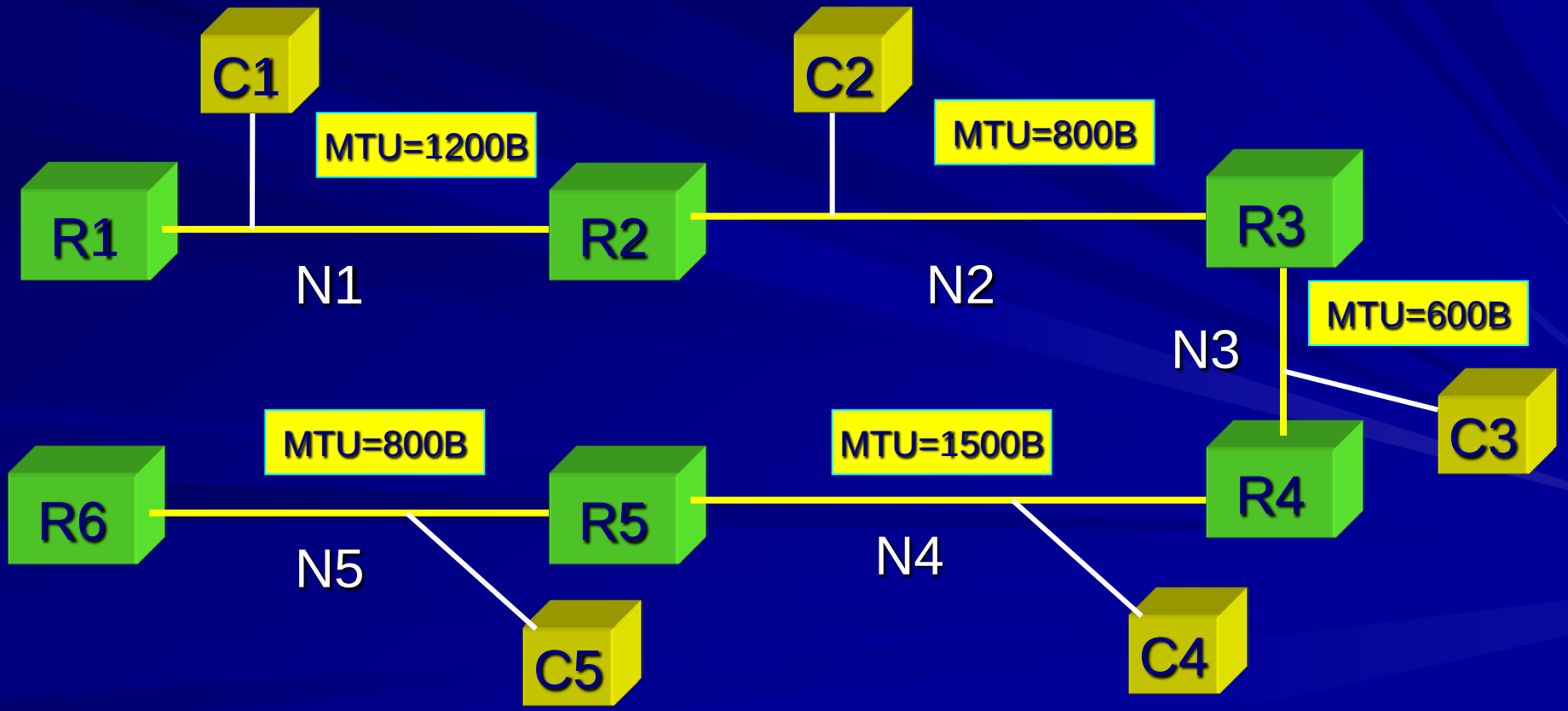

Source host
Net 1
Router 1
Net 2
Router 2
Net 3
Destination host
datagram
header 1
datagram
datagram
header 2
datagram
datagram
header 3
datagram
datagram



(b) 不同层次、不同区间的源地址和目的地址

最大传输单元（MTU）(1)

- 什么是最大传输单元
  - 链路层数据帧支持的最大传输字节数



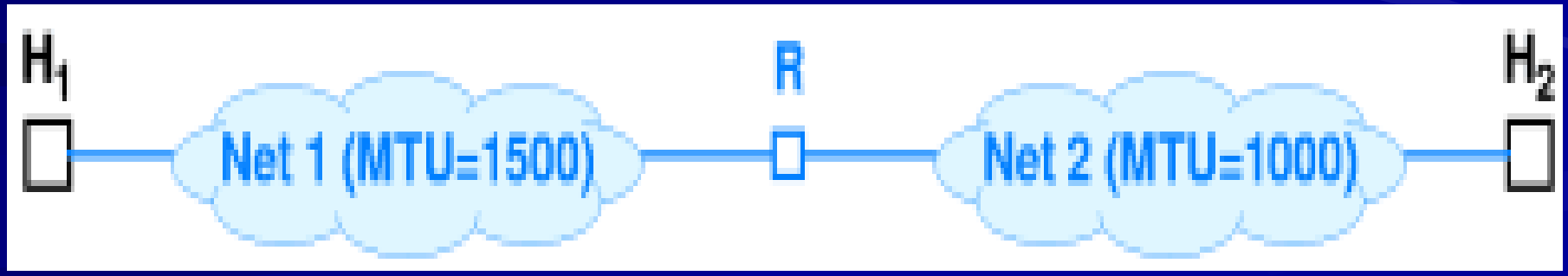
最大传输单元（MTU）(2)

- 路径上最大传输单元对IP报文大小的影响
  - 按照端到端路径上最大MTU作为IP报文长度
    - 1500B
  - 按照端到端路径上最小MTU作为IP报文长度
    - 600B
  - 按照当前链路上MTU作为IP报文长度
    - 1500B
    - 1200B
    - 1000B
    - 800B
    - 600B

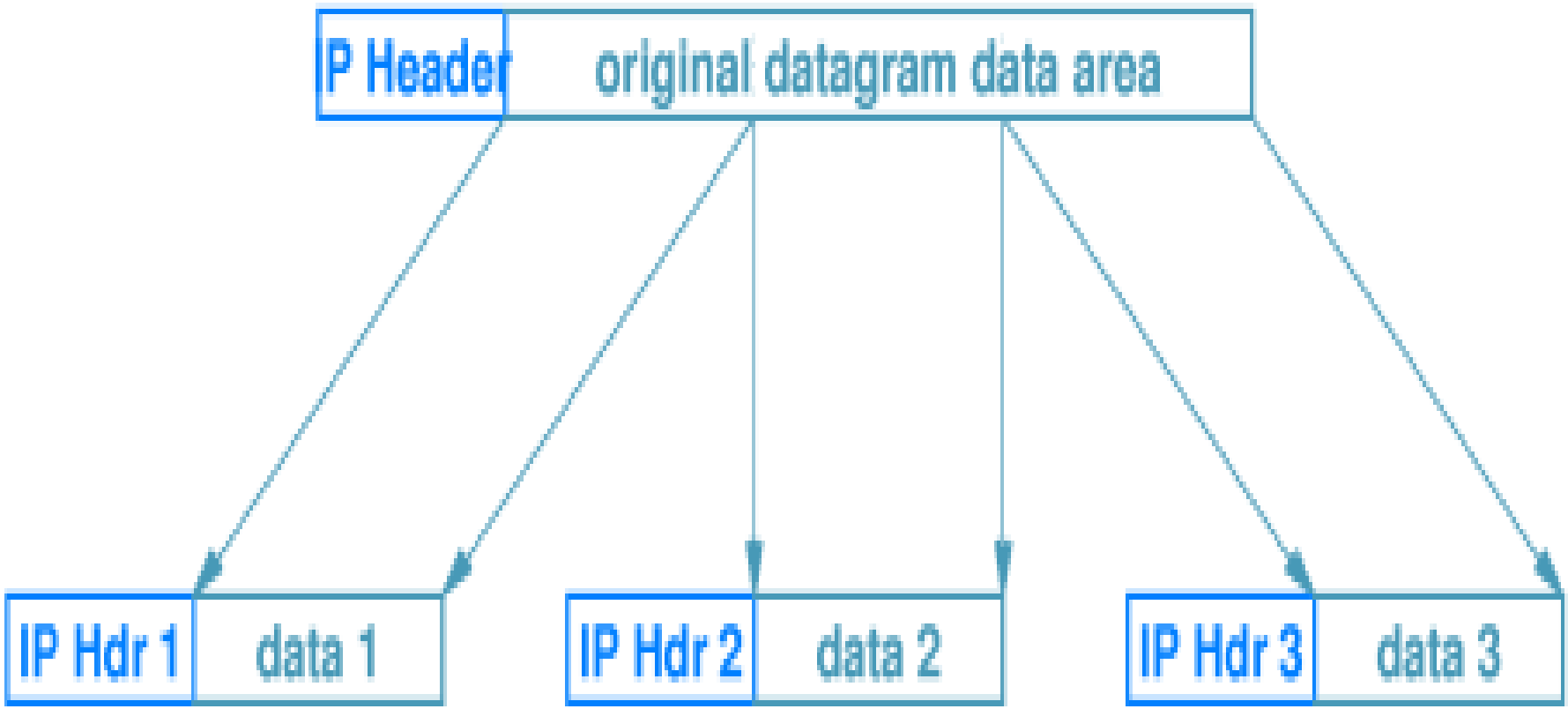
22.5 MTU,Datagram Size, and Encapsulation

- Protocol MTU

Token ring(16Mbps)	17914
Token ring (4Mbps)	4464
FDDI	4352
Ethernet	1500
X.25	576
PPP	296



- The router divides the datagram into smaller pieces called fragments.
- Each fragment uses the IP datagram format.
- sends each fragment independently.

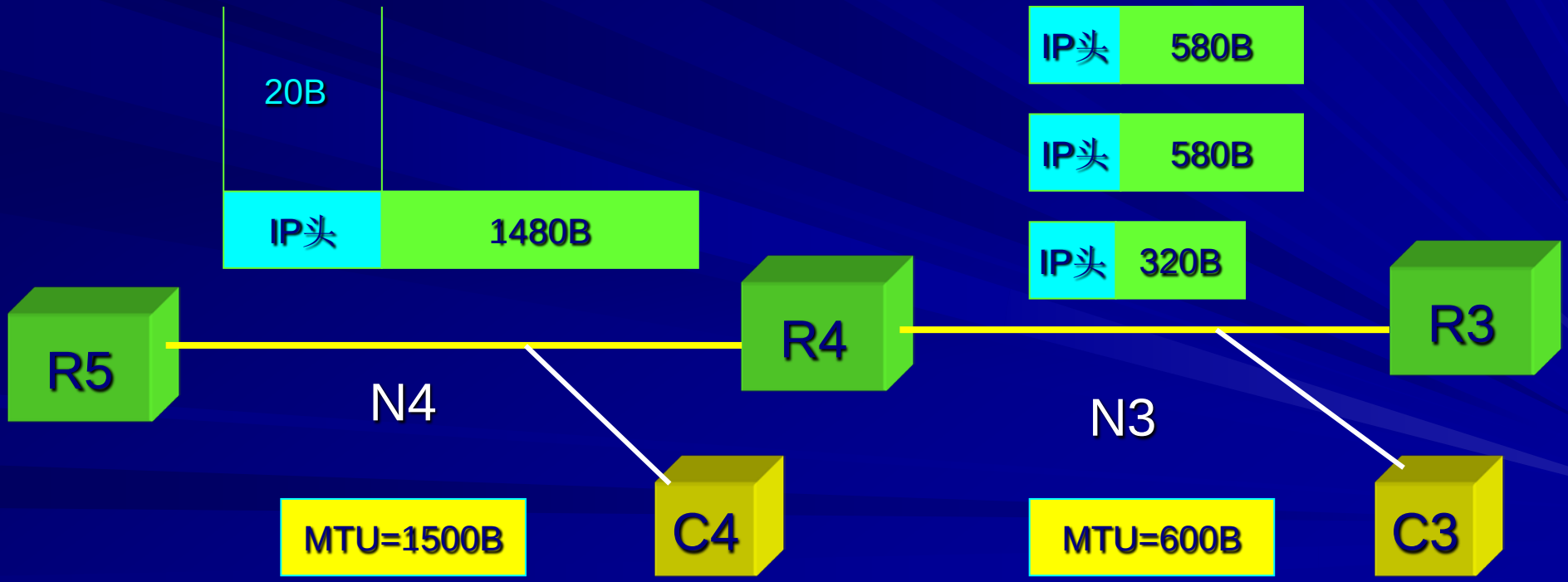


IP报文的分片策略(1)

- IP报文传输的分片策略
  - 在不超过版本本身规定大小的前提下，采用当前最合适的报文长度
  - 在当前链路帧长度小于IP数据报文的时候，将IP报文分成较小的几个分片进行传输
  - 分片原则
    - 各片尽可能大，但是必须能为帧所封装
    - 片大小必须为8的整数倍

IP报文的分片策略(2)

- 分片大小根据MTU、IP报文头长度（基本长度为20字节）计算



分片信息表示

- IP报文中的相关信息
  - IP报文中的ID
    - 始终保持初始ID不变
  - 标志位
    - 如果第一次分片，则修改“是否已经分片”相应位
  - 片偏移量
    - 表示当前分片在初始IP包中有效数据的偏移位置（IP包中以8字节为单位）

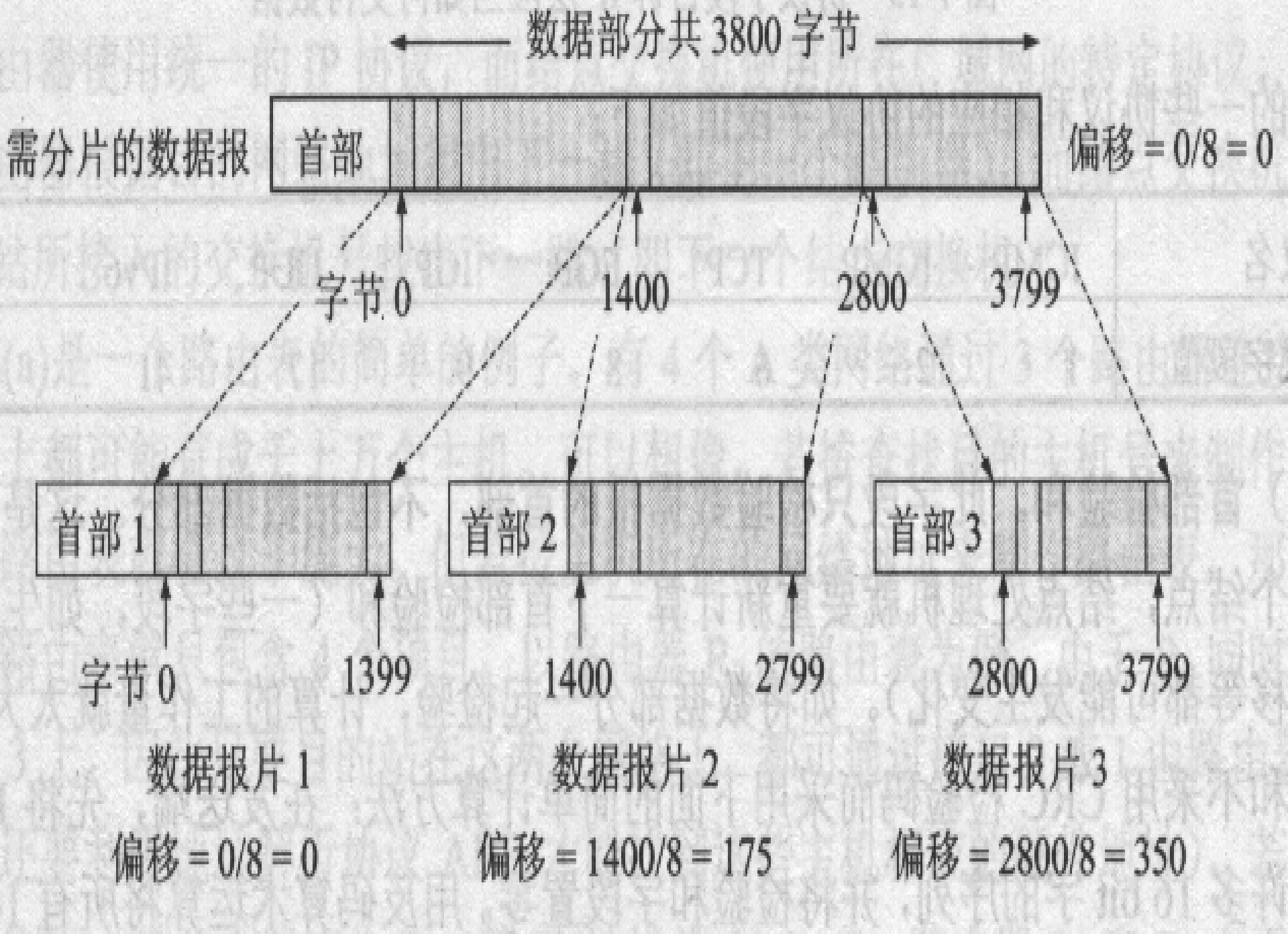

数据部分共3800字节
需分片的数据报
首部
偏移 = 0/8 = 0
字节 0
1400
2800
3799
首部 1
首部 2
首部 3
字节 0
1399
1400
2799
2800
3799
数据报片 1
偏移 = 0/8 = 0
数据报片 2
偏移 = 1400/8 = 175
数据报片 3
偏移 = 2800/8 = 350

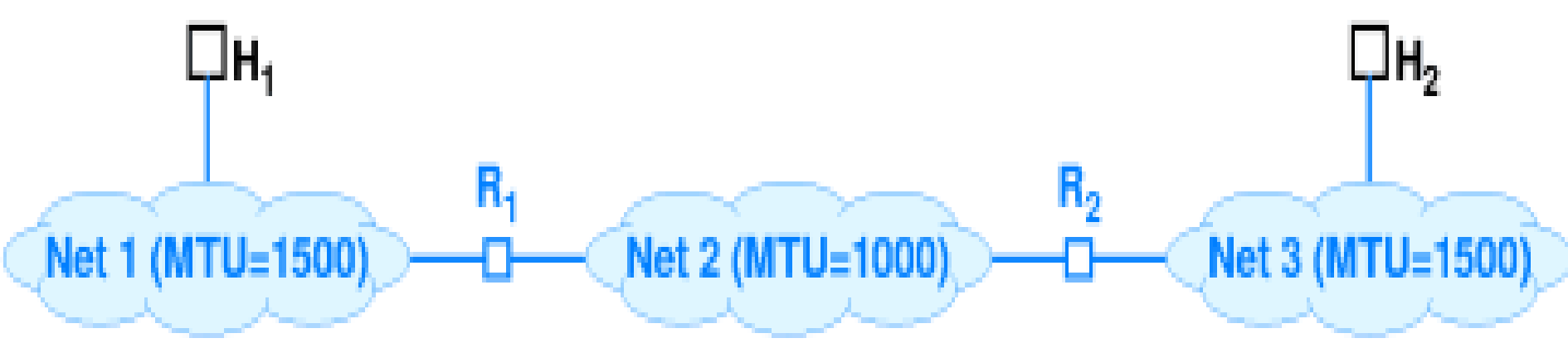
表 7-4　IP 数据报首部中与分片有关的字段中的数值

	总长度	标识	MF	DF	片偏移
原始数据报	4000	12345	0	0	0
数据报片 1	1420	12345	1	0	0
数据报片 2	1420	12345	1	0	175
数据报片 3	1020	12345	0	0	350

- 标志 **MF(more fragment), MF=1** 表示后面还有分片；**MF=0** 表示这是最后一个分片；**DF(don't fragment), DF=1** 表示不允许分片；**DF=0** 表示允许分片。

22.6 Reassembly重组

- The process of creating a copy of the original datagram from fragments is called reassembly.
- IP specifies that the ultimate destination host should reassemble fragments.



22.7 Identifying A Datagram

- A sender places a unique identification number in the IDENTIFICATION field of each outgoing datagram.
- When a router fragments the datagram,the router copies the identification number into each fragment.

IP报文的重组

- 报文重组策略
  - 源端到目标端数据传输过程中可能有多次分片
    - 数据报文大小一定不会大于路径上的最小MTU
  - 所有传输的分片重组在目标端进行
  - 中间路由设备不做分片重组
    - 减少中间节点的数据处理过程

22.8 fragment Loss

- Some fragment are delayed or lost.
- Receiver must save the fragments.
- IP specifies a maximum time to hold fragments.
- Either all fragments arrive and IP reassembles the datagram, or IP discards the complete datagram.
- There is no mechanism for a receiver to tell the sender which fragments have arrived.
- If the sender did retransmit the datagram,routes may be different.

IP报文丢失问题

- IP报文丢失判断
  - 目标端对IP报文分片作重组处理的时候进行丢失判断
  - 对应于源端发出的每一个报文，在收到第一个分片的时候，给出一个等待的有限时间T-out，如果T-out之后还没有收到全部分片，则为超时
  - 任何一个分片丢失或数据出错，则丢弃整个报文

22.9 Fragmenting A fragment

- The fragmentation scheme has been planned carefully to make it possible to further fragment a fragment.

作业

- 从C1向C5发送一个IP报文（报文总长1200B），给出此IP报文在N1、N2、N3、N4、N5中的每个分片的大小及其偏移量。（ IP报头20字节）