
哈爾濱工業大學
HARBIN INSTITUTE OF TECHNOLOGY
立足航天，服务国防，面向国民经济主战场

# 计算机网络

李全龙

# 第6周 课堂教学-网络层（上）

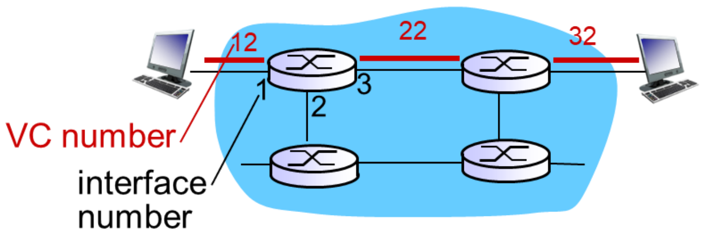
❖ 東广就狭：

- 4.1 网络层服务
- 4.2 转发与路由
- 4.3 虚电路网络与数据报网络
- 4.4 IP协议与IP数据报
- 4.5 IP地址与IP子网
- 4.6 CIDR与路由聚集
- 4.7 DHCP协议

❖质疑辨惑：

1.虚电路的VCID在每段物理链路上为什么取不同值？

## VC forwarding table




*forwarding table in northwest router:*

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

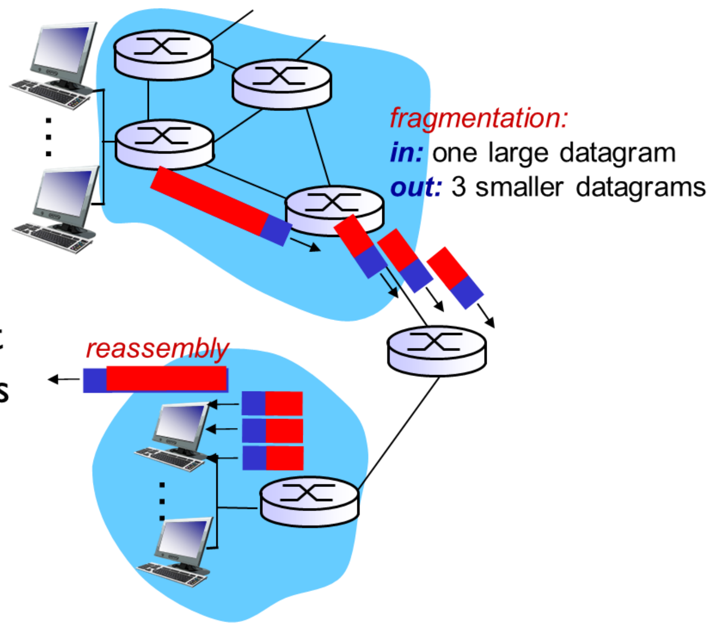
*VC routers maintain connection state information!*

## ❖质疑辨惑：

2.IP分片为什么不在路由器上重组，而在目的主机重组？

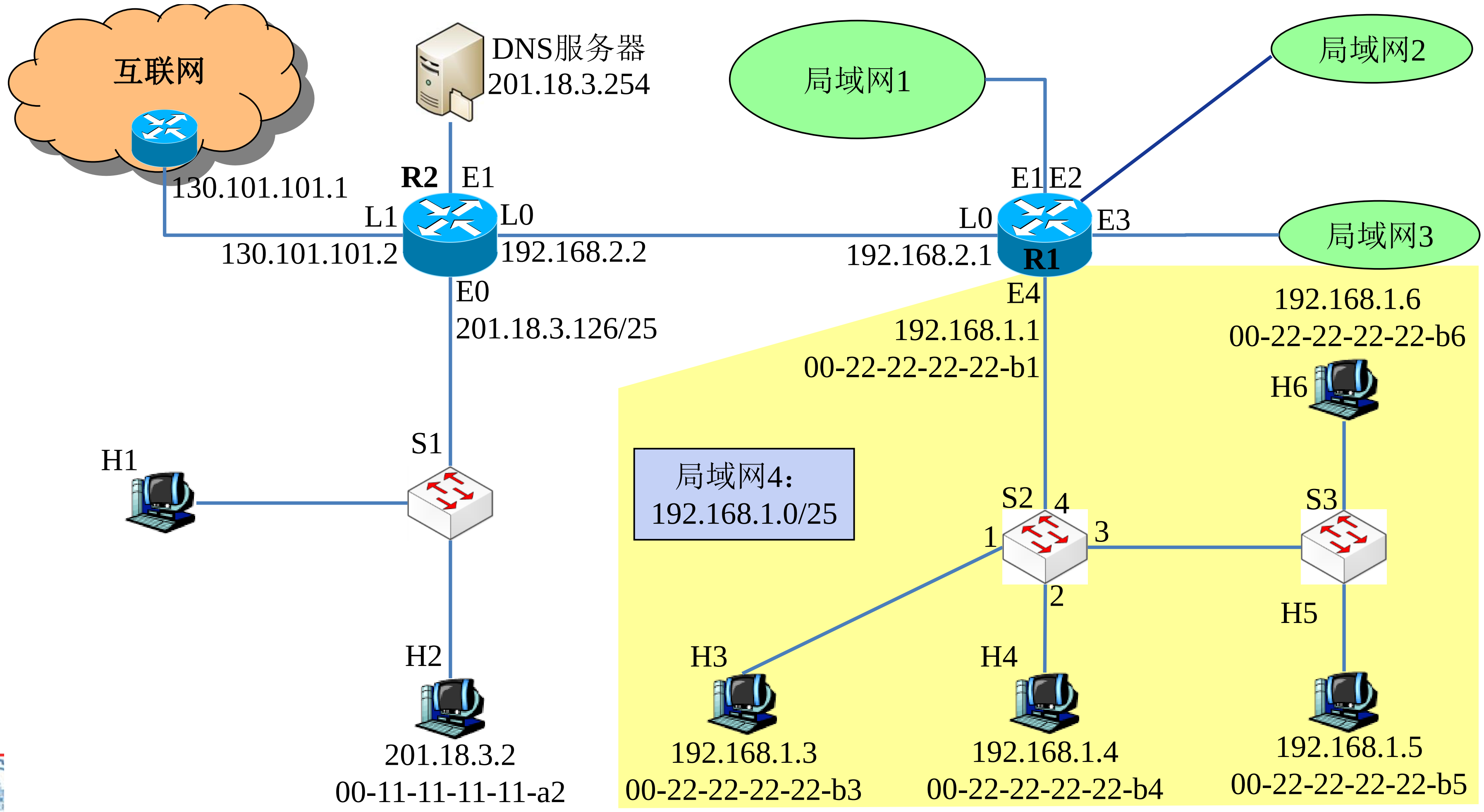
# IP fragmentation, reassembly

- network links have MTU (max.transfer size) - largest possible link-level frame
  - different link types, different MTUs
- large IP datagram divided ("fragmented") within net
  - one datagram becomes several datagrams
  - "reassembled" only at final destination
  - IP header bits used to identify, order related fragments



# 第6周 课堂教学-网络层（上）

❖ 质疑辨惑：3.1 请将192.168.1.0/24剩余IP地址分配给局域网1~3，其中局域网1需要IP地址数不少于60个，局域网2、3需要IP地址数不少于30个。说明局域网1~3的子网地址、广播地址、子网掩码、可分配IP地址数及范围。



# 第6周 课堂教学-网络层（上）

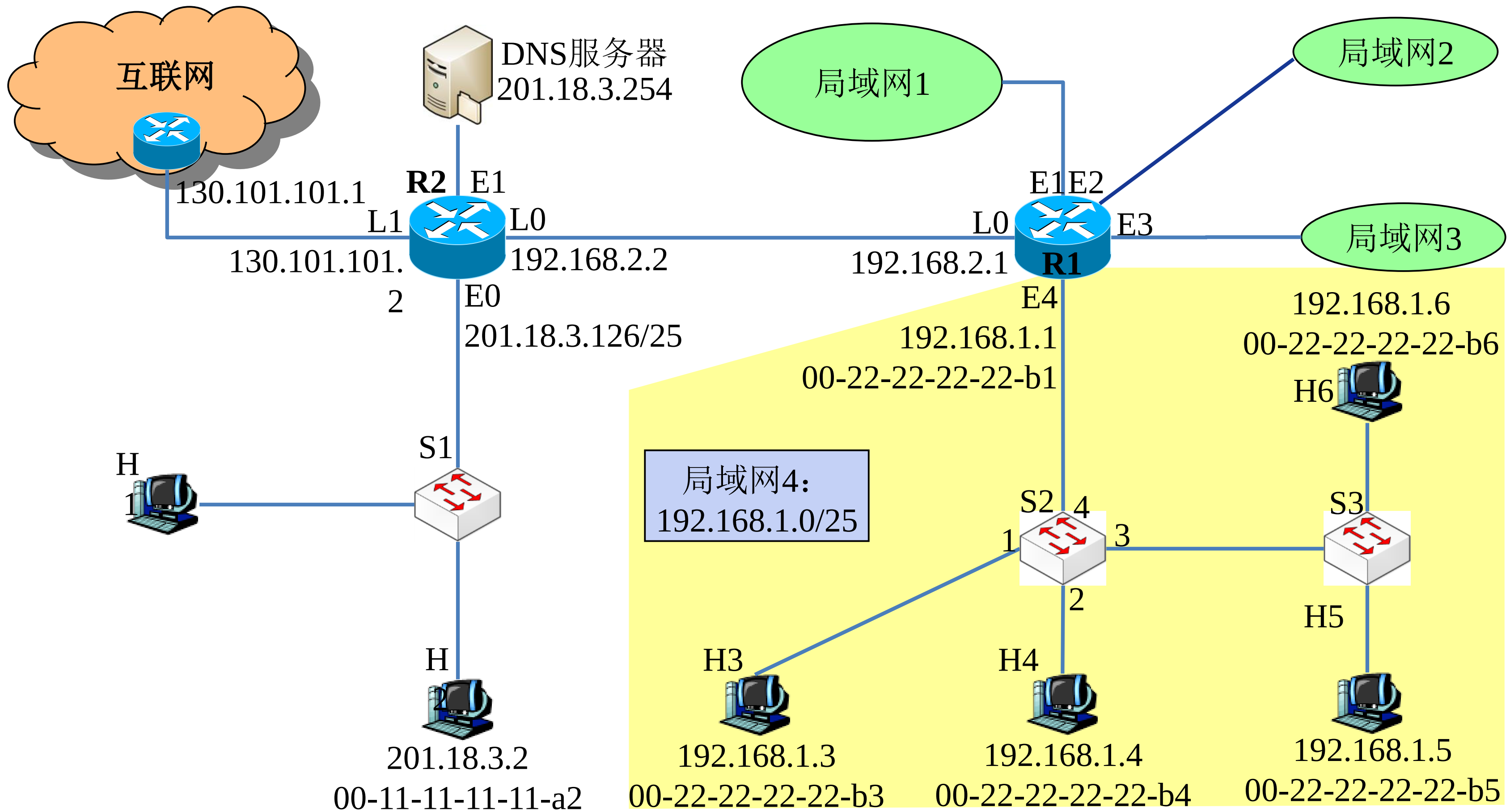
❖ 将剩余IP地址空间（192.168.1.128/25）划分为三个不等长子网：

- 192.168.1.128/26
  - 子网地址：192.168.1.128，广播地址：192.168.1.191，子网掩码：255.255.255.192，可分配IP地址数：62，范围：192.168.1.129~ 192.168.1.190。
- 192.168.1.192/27
  - 子网地址：192.168.1.192，广播地址：192.168.1.223，子网掩码：255.255.255.224，可分配IP地址数：30，范围： 192.168.1.193~ 192.168.1.222。
- 192.168.1.224/27
  - 子网地址：192.168.1.224，广播地址：192.168.1.255，子网掩码：255.255.255.224，可分配IP地址数：30，范围： 192.168.1.225~ 192.168.1.254。

# 第6周 课堂教学-网络层（上）

❖ 质疑辨惑： 3.2 给出R1的路由表。路由表结构如下：

| 目的网络 | 子网掩码 | 下一跳（IP地址） | 接口 |
|---|---|---|---|

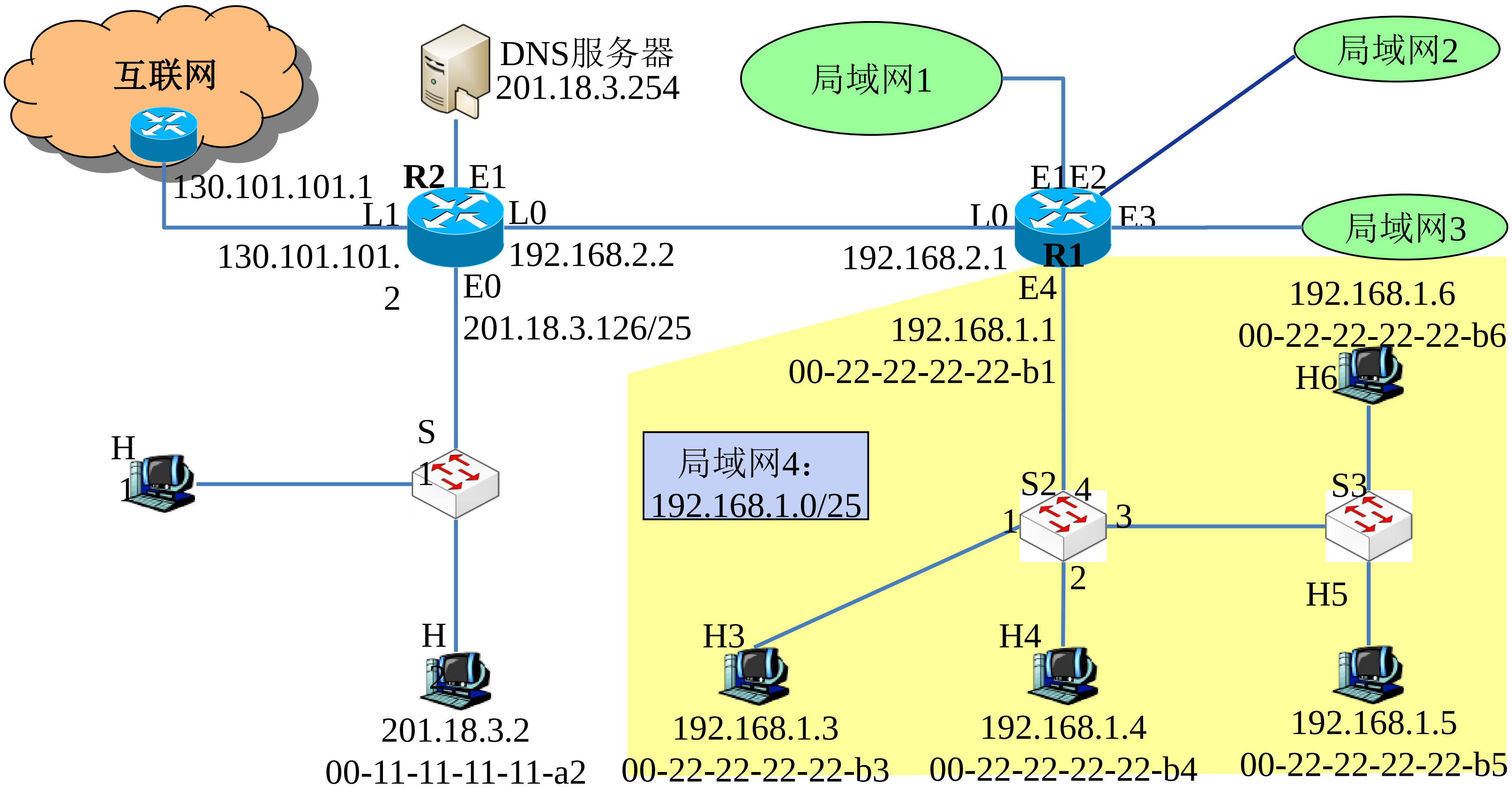


## R1的路由表：

| 目的网络 | 子网掩码 | 下一跳（IP地址） | 接口 |
|---|---|---|---|
| 192.168.1.128 | 255.255.255.192 | — | E1 |
| 192.168.1.192 | 255.255.255.224 | — | E2 |
| 192.168.1.224 | 255.255.255.224 | — | E3 |
| 192.168.1.0 | 255.255.255.128 | — | E4 |
| 201.18.3.0 | 255.255.255.128 | 192.168.2.2 | L0 |
| 201.18.3.254 | 255.255.255.255 | 192.168.2.2 | L0 |
| 0.0.0.0 | 0.0.0.0 | 192.168.2.2 | L0 |

# 第6周 课堂教学-网络层（上）

❖ 质疑辨惑：3.3 基于路由聚集给出路由器R2中关于到达局域网1~4的路由？为什么要路由聚集？什么条件下可以进行路由聚集?

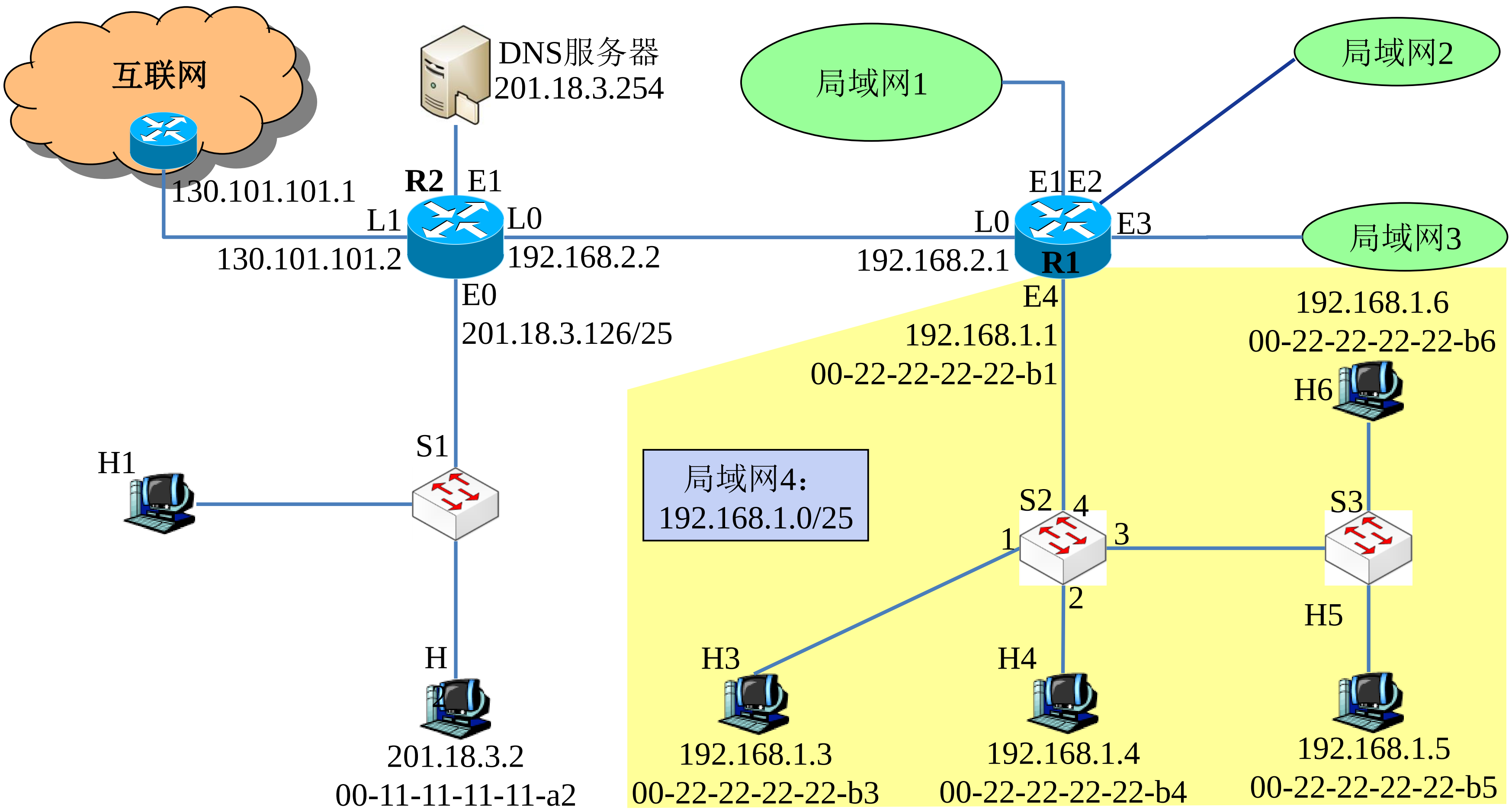


## R2的路由表（到达局域网1~4的路由）：

| 目的网络 | 子网掩码 | 下一跳（IP地址） | 接口 |
| --- | --- | --- | --- |
| 192.168.1.0 | 255.255.255.0 | 192.168.2.1 | L0 |
| …… | …… | …… | …… |

- 路由聚合条件：
  - 具有共同的网络前缀
  - 具有相同的下一跳
  - 具有相同的接口
  - 不产生路由“黑洞”
    - 最长前缀匹配优先

❖ 质疑辨惑： 3.4 请为主机H1配置IP地址信息，包括IP地址、子网掩码、默认网关以及域名服务器。

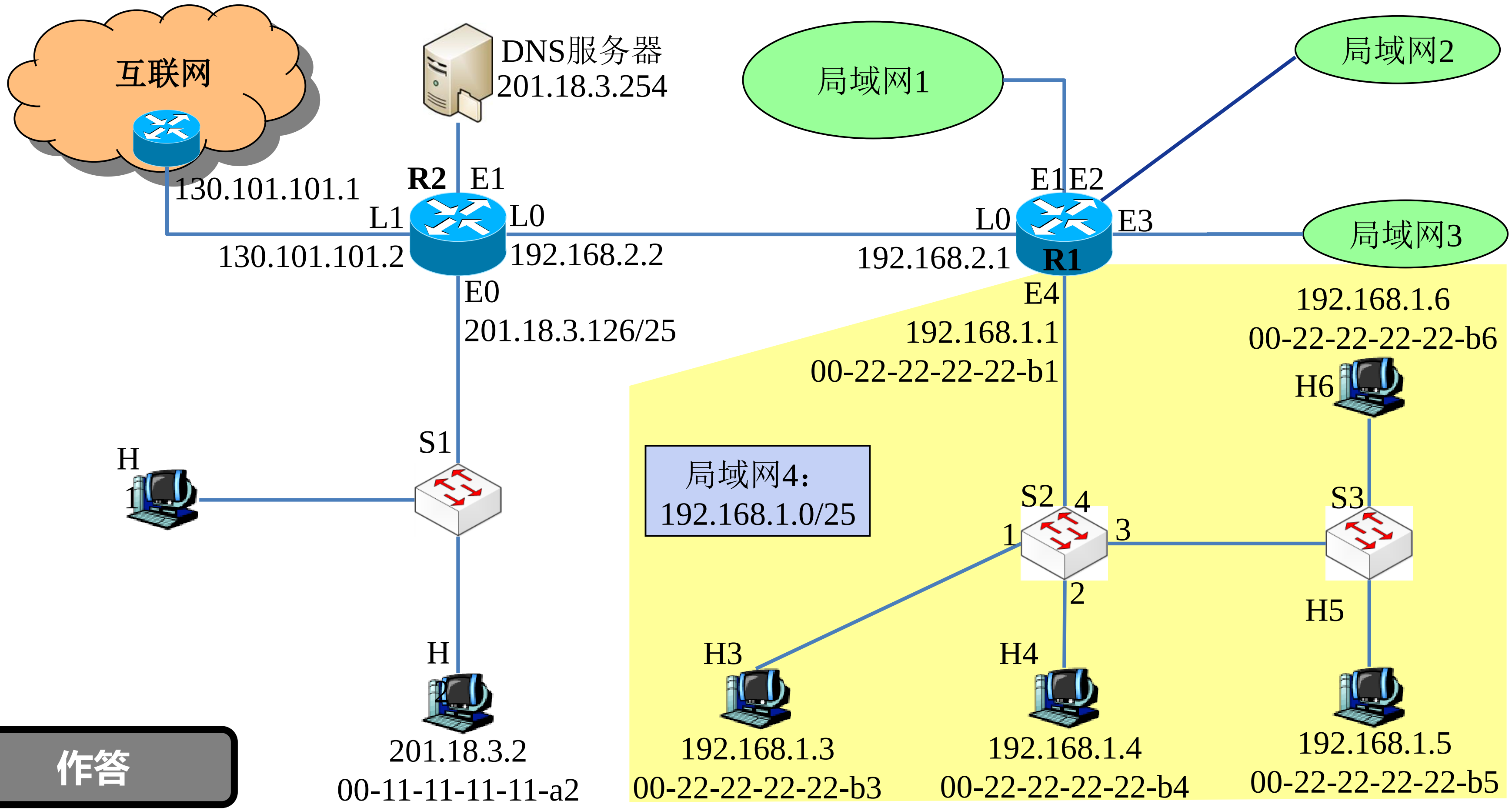


- H1的IP地址：201.18.3.1
- 子网掩码：255.255.255.128
- 默认网关：201.18.3.126
- 域名服务器：201.18.3.254

主观题 10分

设置

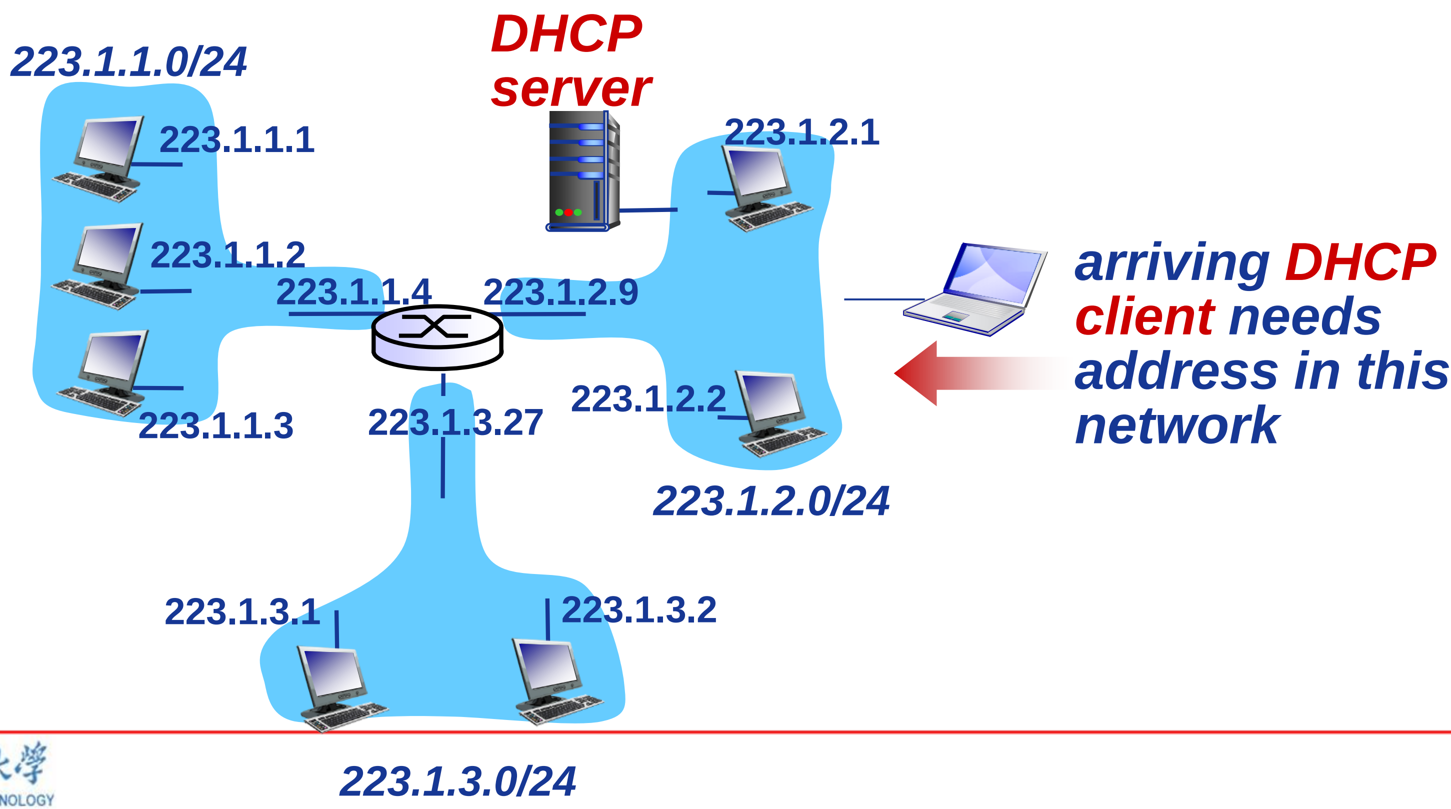
❖ 质疑辨惑：3.5 若主机H2的子网掩码被配置为255.255.255.0，DNS配置为201.18.3.254，则H2能否成功访问www.sina.com.cn？为什么？




作答

❖ 质疑辨惑：

4.通过DHCP动态分配IP地址过程中需要交换哪些DHCP报文？这些报文直接封装到哪个协议的数据包中？封装这些报文的IP数据报的目的IP地址是什么？为什么？

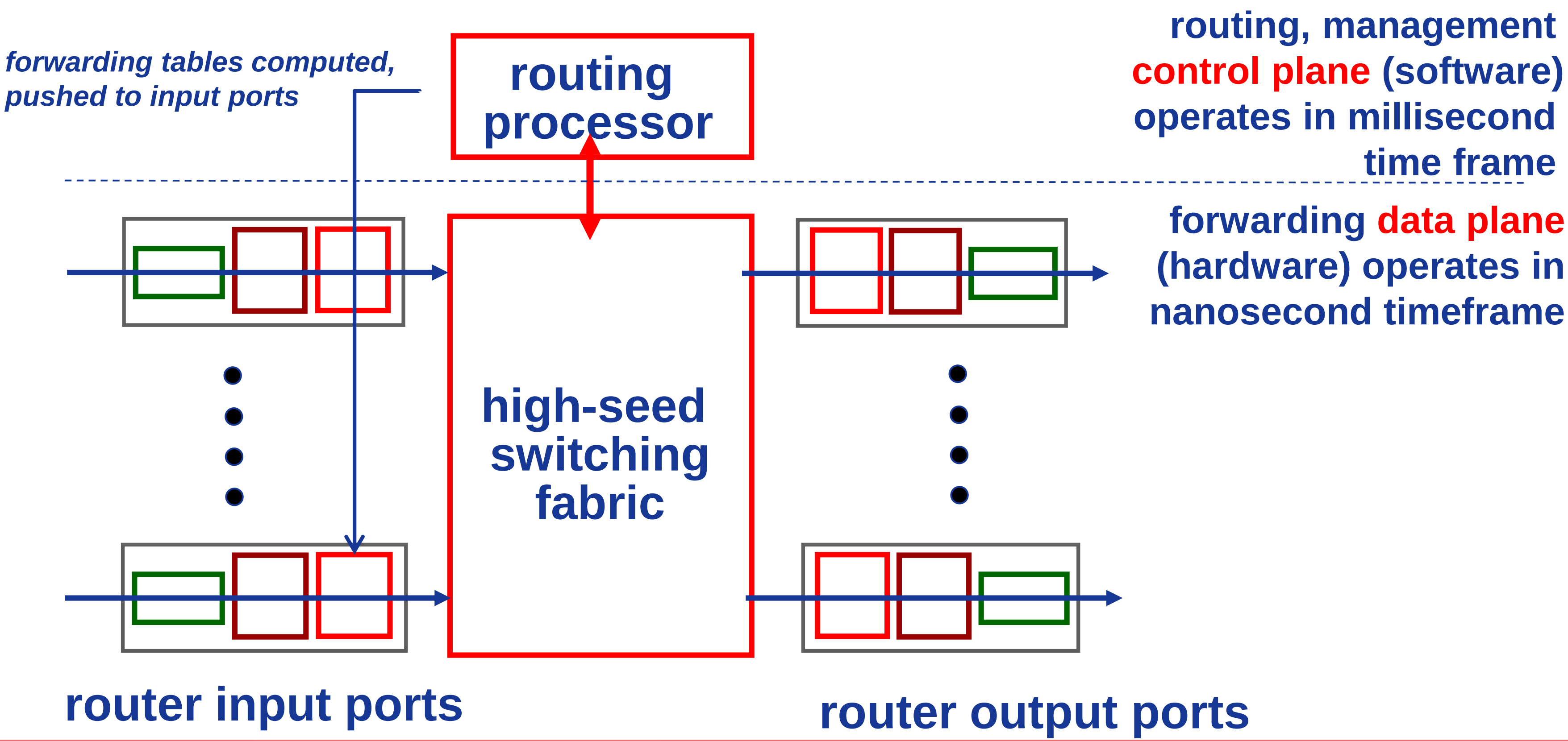


- 开疆拓土：
  - 路由器体系结构

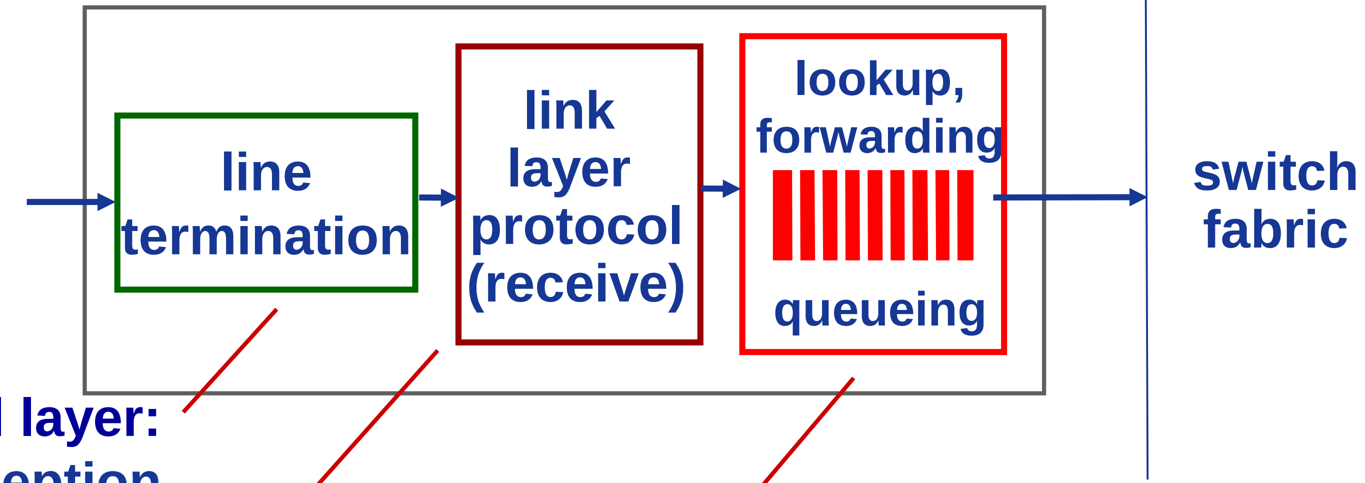
# Router architecture overview

two key router functions:

- run routing algorithms/protocol (RIP, OSPF, BGP)
- *forwarding* datagrams from incoming to outgoing link



# Input port functions



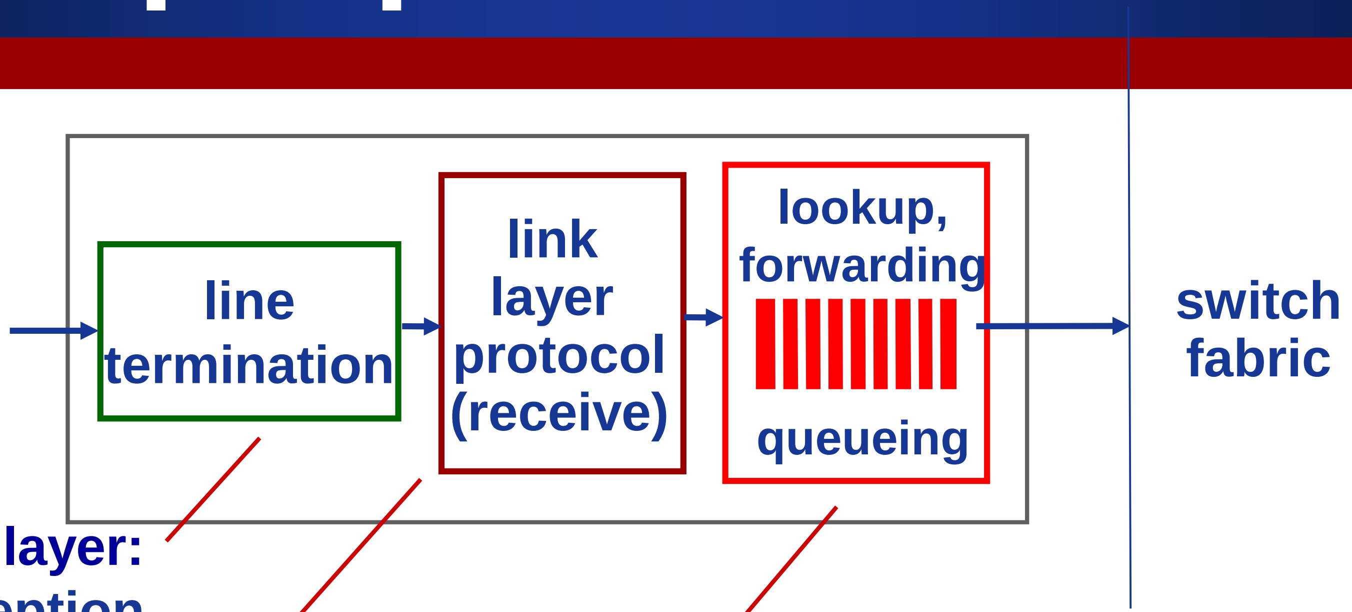

**physical layer:**
**bit-level reception**

**data link layer:**
**e.g., Ethernet**
**see chapter 5**

decentralized switching*:*

- given datagram dest., lookup output port using forwarding table in input port memory *("match plus action")*
- goal: complete input port processing at 'line speed'
- queuing: if datagrams arrive faster than forwarding rate into switch fabric

# Input port functions




**physical layer:**
**bit-level reception**

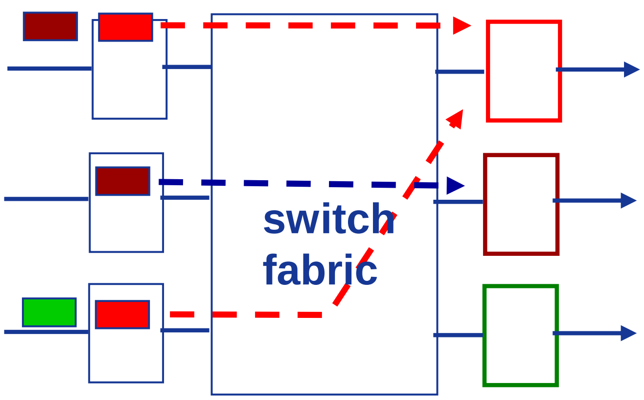
**data link layer:**
**e.g., Ethernet**
**see chapter 5**

decentralized switching*:*

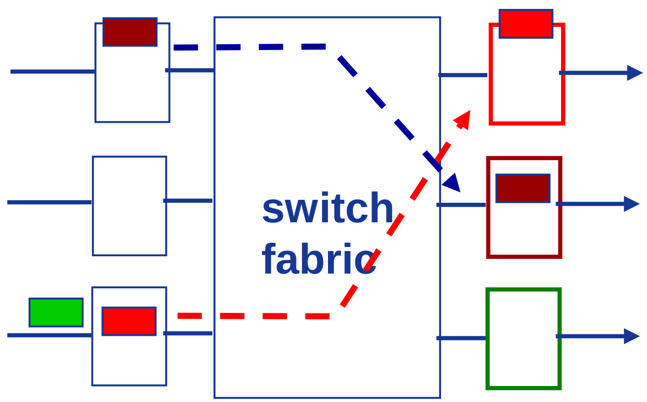
- using header field values, lookup output port using forwarding table in input port memory *("match plus action")*
- *destination-based forwarding:* forward based only on destination IP address (traditional)
- *generalized forwarding:* forward based on any set of header field values

# Input port queuing

- fabric slower than input ports combined -> queueing may occur at input queues
  - *queueing delay and loss due to input buffer overflow!*
- Head-of-the-Line (HOL) blocking: queued datagram at front of queue prevents others in queue from moving forward



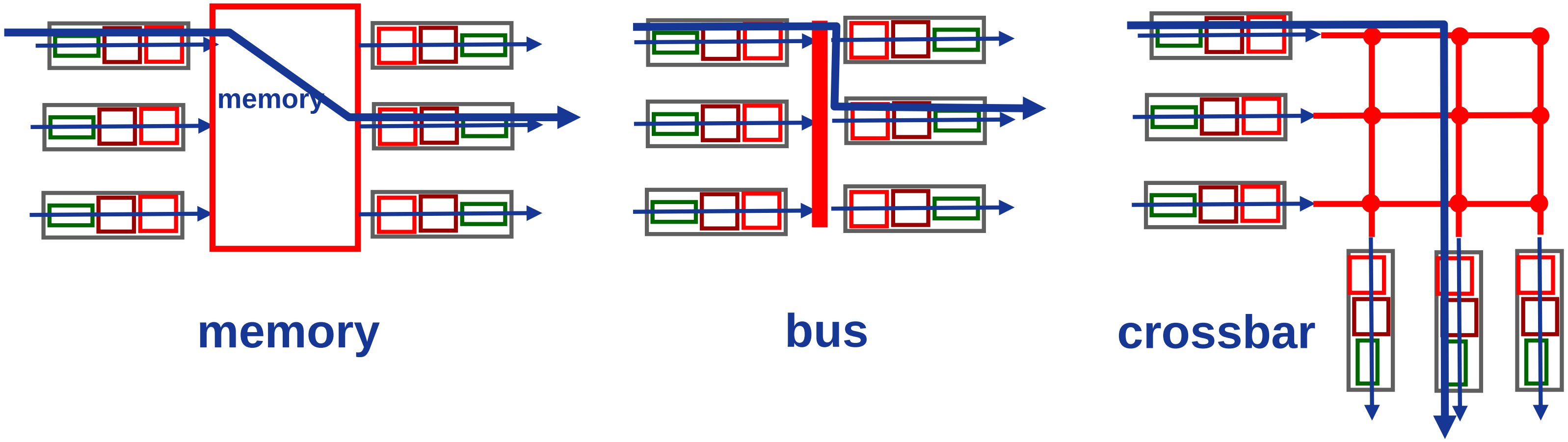

**output port contention: only one red datagram can be transferred.**
***lower red packet is blocked***




**one packet time later: green packet experiences HOL blocking**

# Switching fabrics

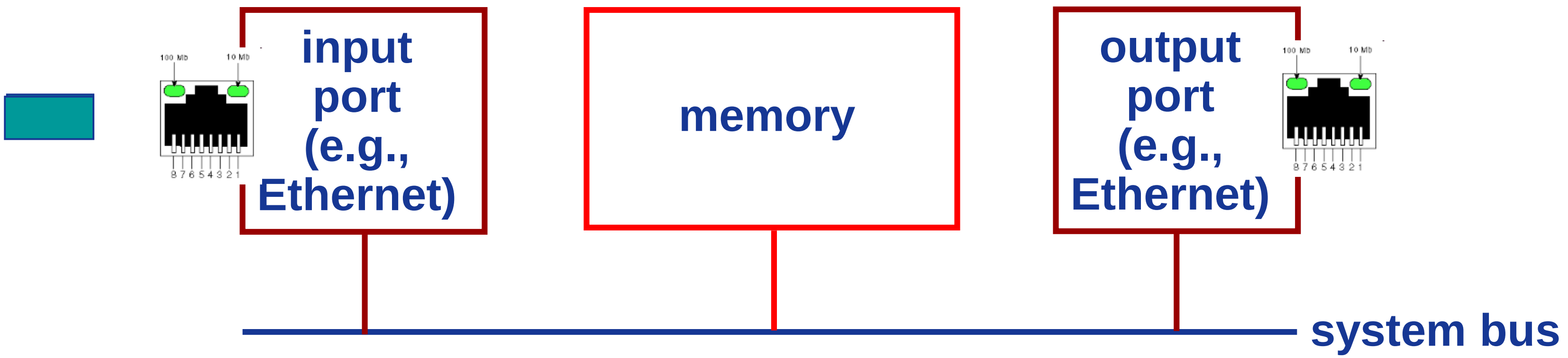
- transfer packet from input buffer to appropriate output buffer
- switching rate: rate at which packets can be transfer from inputs to outputs
  - often measured as multiple of input/output line rate
  - N inputs: switching rate N times line rate desirable
- three types of switching fabrics



# Switching via memory

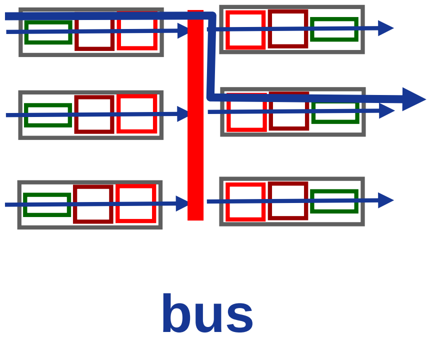
*first generation routers:*

- 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)



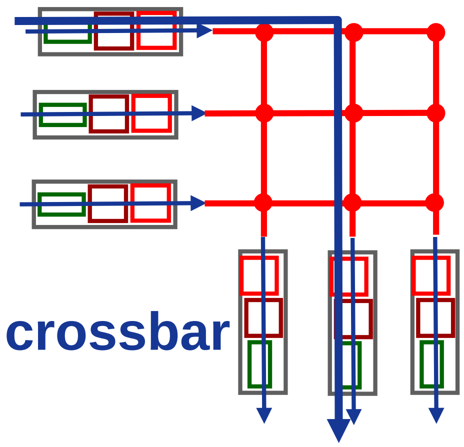
# Switching via a bus

- 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

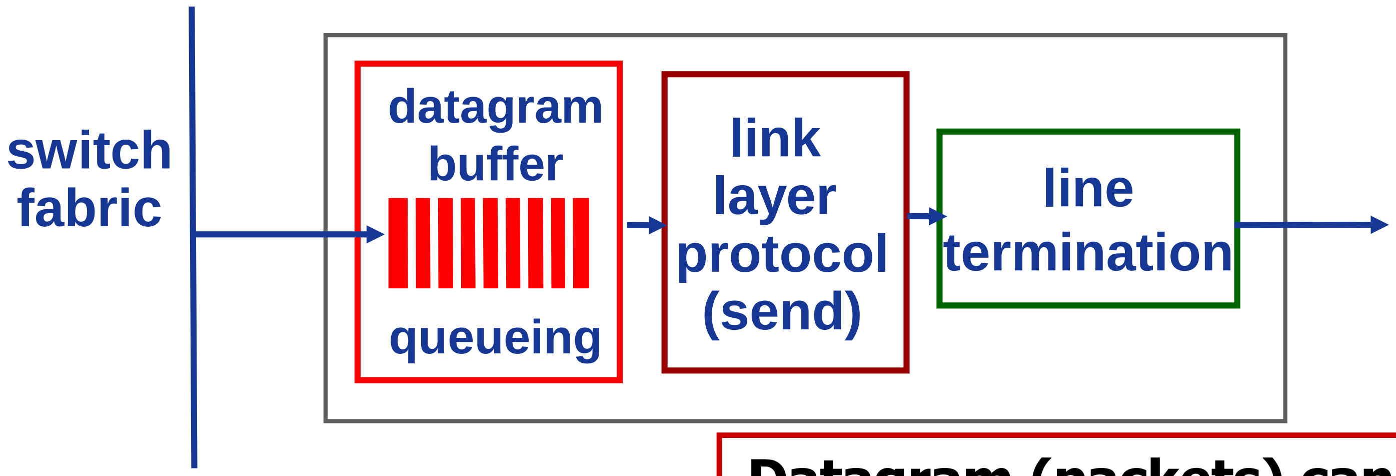


# Switching via interconnection network

- 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 ports

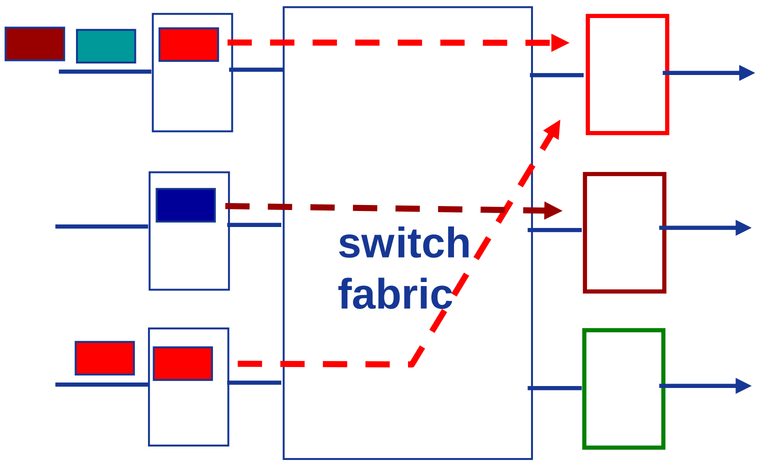



**Datagram (packets) can be lost due to congestion, lack of buffers**

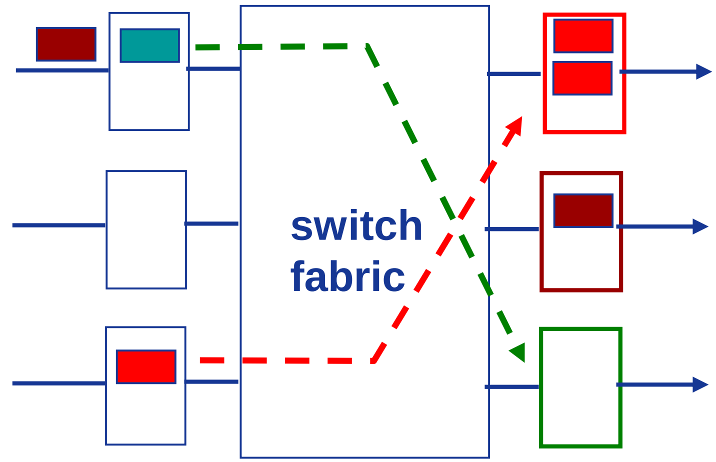
- *buffering* required when datagrams arrive from fabric faster than the transmission rate
- *scheduling discipline* chooses among queued datagrams for transmission

**Priority scheduling – who gets best performance, network neutrality**

# Output port queueing




**at *t*, packets more from input to output**




**one packet time later**

- buffering when arrival rate via switch exceeds output line speed
- *queueing (delay) and loss due to output port buffer overflow!*

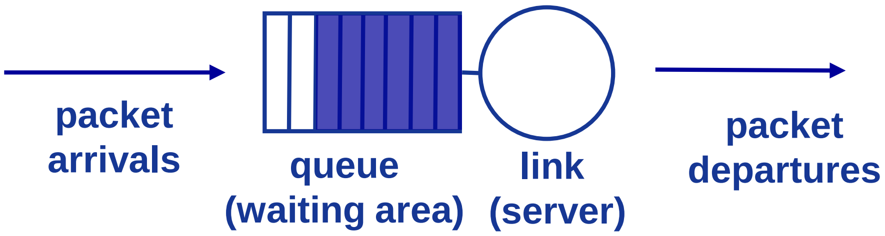
# How much buffering?

- 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

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

# Scheduling mechanisms

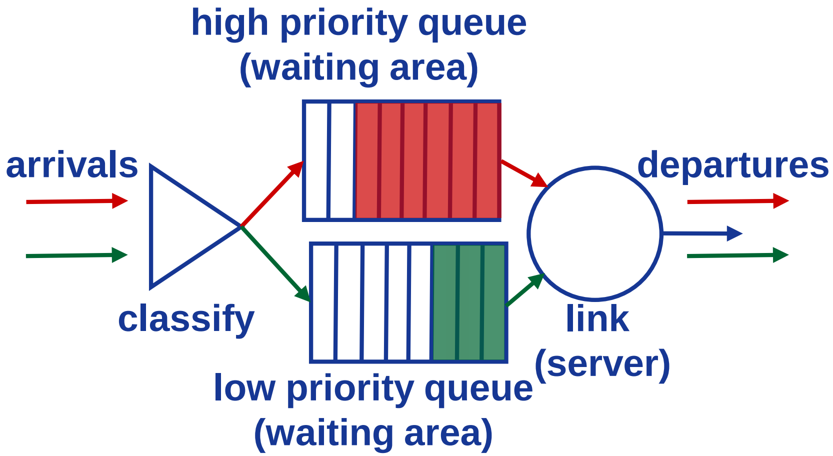
- *scheduling:* choose next packet to send on link
- *FIFO (first in first out) scheduling:* send in order of arrival to queue
  - real-world example?
  - *discard policy:* if packet arrives to full queue: who to discard?
    - *tail drop:* drop arriving packet
    - *priority:* drop/remove on priority basis
    - *random:* drop/remove randomly

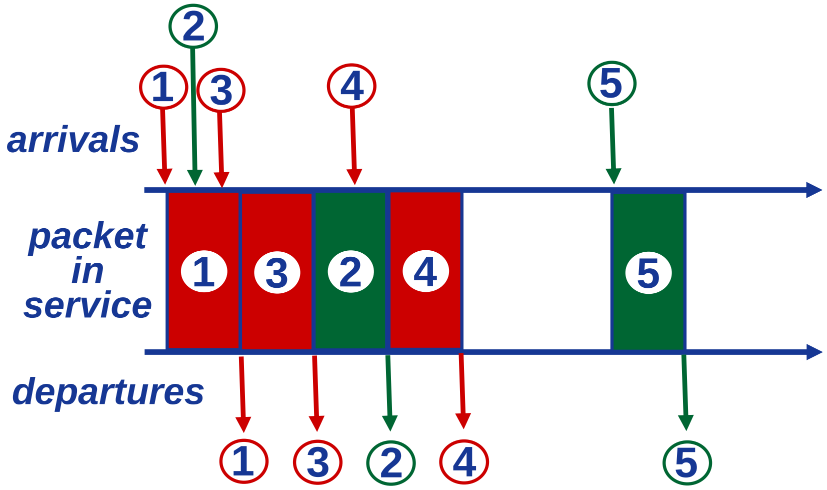


# Scheduling policies: priority

*priority scheduling:* send highest priority queued packet

- multiple *classes*, with different priorities
  - class may depend on marking or other header info, e.g. IP source/dest, port numbers, etc.
  - real world example?

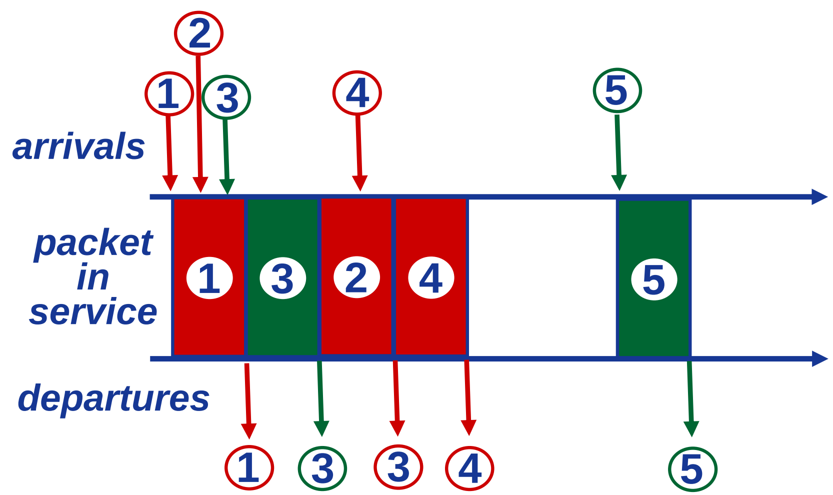


# Scheduling policies: still more

*Round Robin (RR) scheduling:*

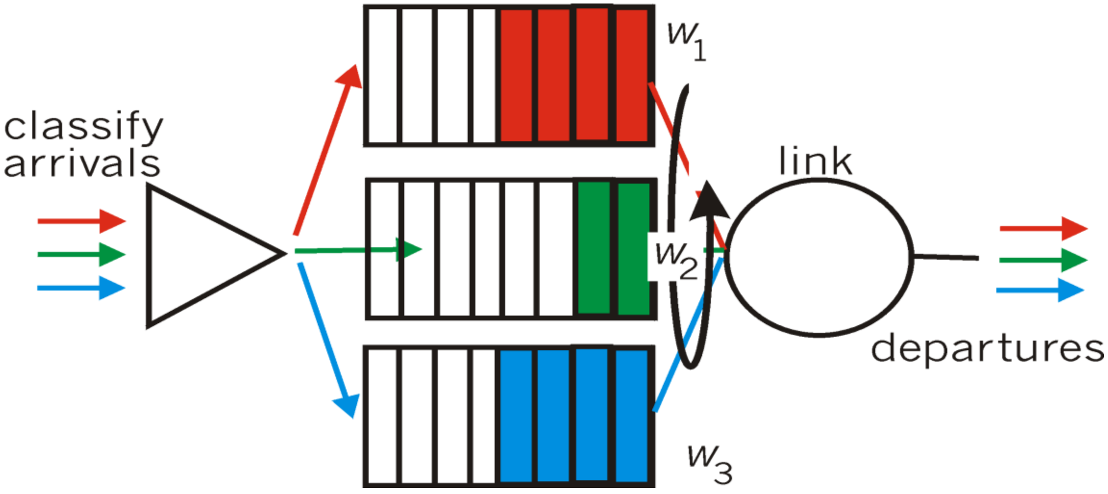
- multiple classes
- cyclically scan class queues, sending one complete packet from each class (if available)
- real world example?



# Scheduling policies: still more

*Weighted Fair Queuing (WFQ):*

- generalized Round Robin
- each class gets weighted amount of service in each cycle
- real-world example?



# 第6周 课堂教学-网络层（上）

## 解疑释惑：

- 1.虚电路网络有什么特点？数据报网络有什么特点？两者有什么共同点？
- 2.IP数据报首部长度字段占几位？其单位是什么？
- 3.IP数据报首部的片偏移量的单位是什么？为什么？
- 4.IP分组首部的标志位(flags)有几位？分别是什么作用？
- 5.IP分组分片时涉及到哪些字段？这些字段如何取值？
- 6.什么是子网掩码？如何取值？作用是什么？
- 7.什么是默认网关？作用是什么？
- 8.端系统如何探测到达另一端系统的路径的最小MTU？
- 9.如何将一个IP网络划分为指定数量的子网？
- 10.如何确定一个子网的可分配IP地址数？

- 演武修文：
  - 课堂测验

某主机的IP地址为180.80.77.55，子网掩码为255.255.252.0。若该主机向其所在子网发送广播分组，则目的地址可以是

A 180.80.76.0

B 180.80.76.255

C 180.80.77.255

D 180.80.79.255

提交

# 在子网192.168.4.0/30中，能够接收目的地址为192.168.4.3的IP分组的最大主机数是

A 0

B 1

C 2

D 4

提交

若将101.200.16.0/20划分为5个子网，则可能的最小子网的可分配IP地址数是________。

A 126

B 254

C 510

D 1022

提交

# 第7周 课堂教学-网络层（下）

- 束广就狭：（30分钟）**第6组报告总结**
  - 总结ICMP协议，NAT，IPv6简介，路由基本原理与算法，典型路由算法（链路状态算法与距离矢量算法），层次化路由。
- 质疑辨惑：（50分钟）
  - 1.什么是路由聚集？为什么要路由聚集？什么条件下可以进行路由聚集？
  - 2.ICMP协议的作用是什么？
  - 3.如果两个均使用私有IP地址的主机需要进行P2P通信？可能会遇到什么问题？如何解决？
  - 4.距离向量路由算法可能会产生什么问题？如何消解该问题？
  - ……
- 解疑释惑：（10分钟）
  - 解答疑问
- 演武修文：（10分钟）
  - 课堂测验
  - 讲解


哈爾濱工業大學
HARBIN INSTITUTE OF TECHNOLOGY
立足航天，服务国防，面向国民经济主战场

# 谢 谢！