



西北師範大學
NORTHWEST NORMAL UNIVERSITY

计算机网络

第四章 网络层 (5)

陈旺虎

chenwh@nwnu.edu.cn



Review



- 路由器转发分组的流程
- 划分子网
- CIDR

习题1

- 对一个局域网进行子网划分
 - ▣ 三个子网分别包含1080、790和28台计算机
 - ▣ 分配给该局域网一个B类地址169.78.0.0
 - ▣ 请说明IP分配方案

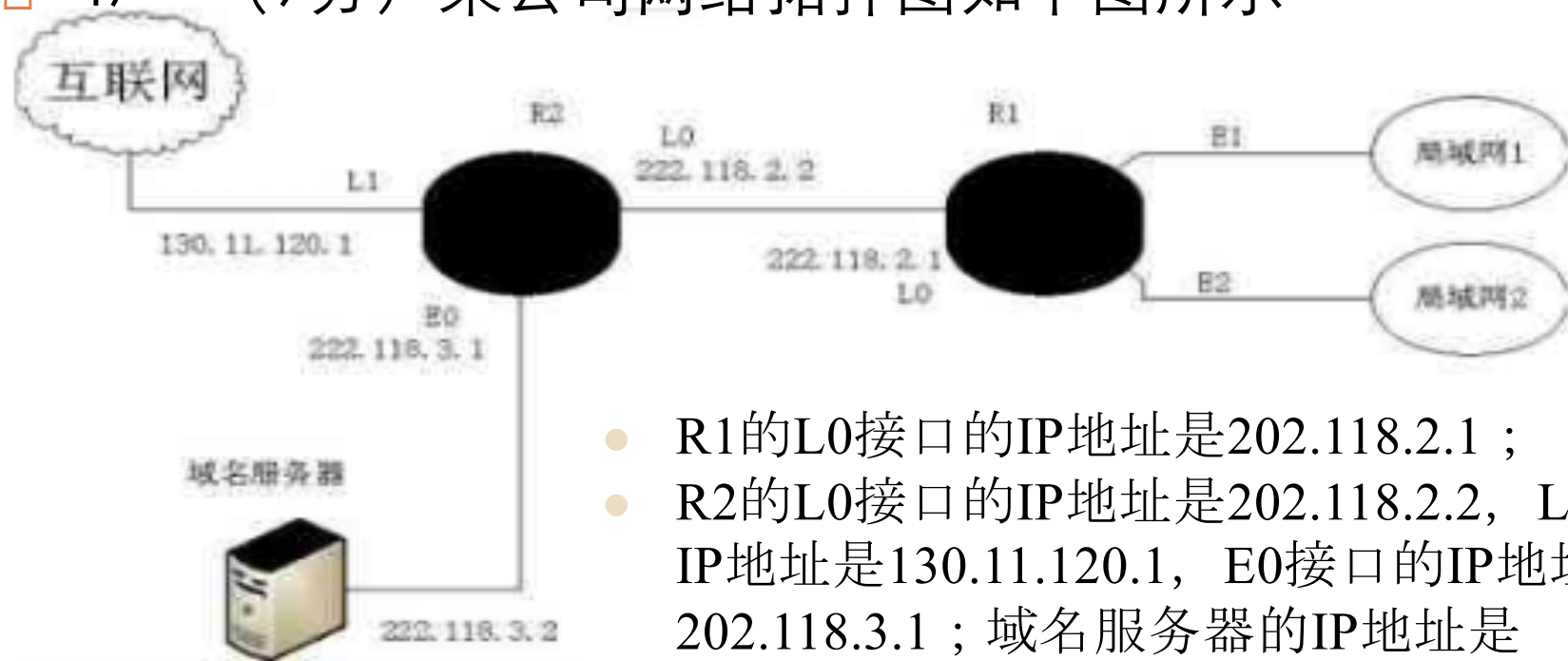
子网号	子网掩码	子网地址	最小IP地址	最大IP地址

- 
- 37、某网络的IP地址为192.168.5.0/24，采用子网划分，子网掩码为255.255.255.248，则该网络的最大子网个数、每个子网内的最大可分配地址个数为（ ）
 - A: 32, 8
 - B: 32, 6
 - C: 8, 32
 - D: 8, 30

解答

- 可知，子网号为5位
- 可以表示 $2^5=32$ 个子网
- 主机号为3位，除去全0和全1的情况可以表示6个主机地址
- 答案： B

47 . （9分）某公司网络拓扑图如下图所示



- R1 的 L0 接口的 IP 地址是 202.118.2.1 ；
- R2 的 L0 接口的 IP 地址是 202.118.2.2， L1 接口的 IP 地址是 130.11.120.1， E0 接口的 IP 地址是 202.118.3.1 ；域名服务器的 IP 地址是 202.118.3.2

R1 和 R2 的路由表结构为：

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

- 将 IP 地址空间 202.118.1.0/24 划分为两个子网，分配给局域网 1、局域网 2，每个局域网分配的地址数不少于 120 个，请给出子网划分结果。
- 给出 R1 和 R2 的路由表。

解答

- 每个子网120台主机，因此，主机号的位数 x 应满足：

- $x < 8$

- $2^x > 120$

- 因此，可得到 $x=7$ ，则，子网掩为

11111111 11111111 11111111 10000000

即：**255.255.255.128**

- 所以划分的两个网段是：**202.118.1.0/25**与**202.118.1.128/25**。

- 
- 3. 某一网络的一台主机产生了一个IP数据报，头部长度是20字节，数据部分长度为2000字节。
 - ▣ 该数据报需要经过两个网络到达目的主机，这两个网络的最大传输单元MTU分别为1500字节和576字节。
 - ▣ 请问IP数据报到达目的主机时，共分为了多少个小IP报文？长度分别是多少？

- (1) 经过网络1: $1480+20$, $520+20$
- (2) 经过网络2: $556+20$ $556+20$ $368+20$
 $520+20$

- 注意整个IP报文作为数据链路层的数据部分，以太网的MTU为1500
- 分片可以在终点组装，也可以没通过一个网络组装

作业



□ 10, 17, 20, 21, 22, 26, 28

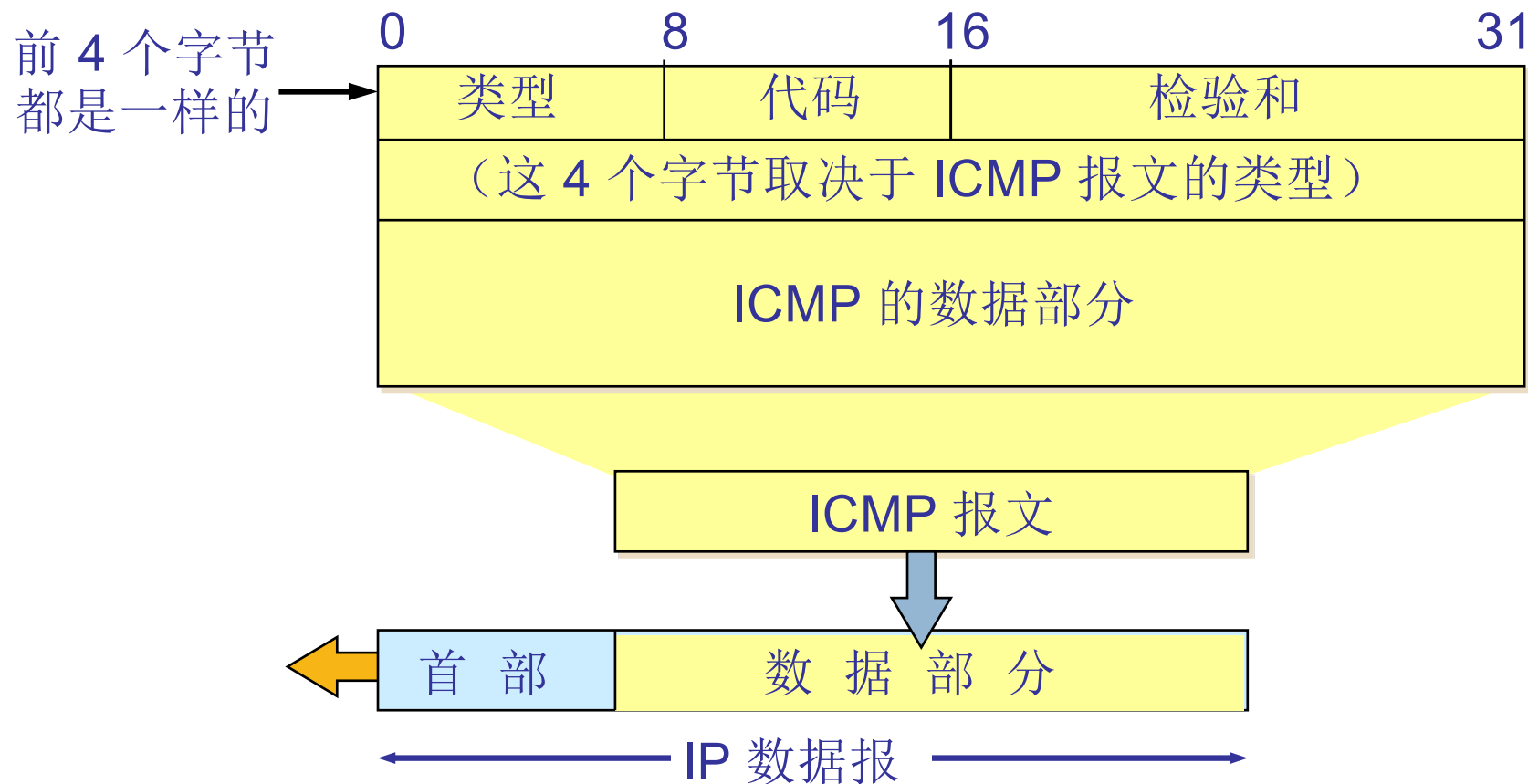
4.4 网际控制报文协议 ICMP

- IP层的服务定位
- IP数据报的校验和错误检测
 - ▣ 每个IP数据报的头部都有一个校验和(checksum)字段，它负责对整个IP报文头进行检验
 - ▣ 路由器在每次转发一个IP报文之前都要重新计算一个新的校验和。（原因？）
- IP层发现校验和错时的处理：数据报丢弃
 - ▣ 因为，接收者无法相信数据报头部中的任何域，包括：源地址，目标地址。
- 但是，某些差错是可以被报告的
 - ▣ 例如，网中的一些物理路径出错，导致网络无法连通等

ICMP的功能

- 为了提高 IP 数据报交付成功的机会，在网际层使用了网际控制报文协议ICMP
 - ▣ ICMP允许主机或路由器报告差错情况和提供有关异常情况的报告。
- ICMP是一个差错报告协议
 - ▣ 检查并报告一些基本的差错，让一个路由器向其它路由器或主机发送差错或控制报文
 - ▣ ICMP在两台机器上的Internet协议软件之间提供了一种通信方式
 - ▣ ICMP协议对IP的执行是必要的

ICMP 报文的格式



- ❑ IP和ICMP相互依赖，IP使用ICMP发送差错报文，ICMP利用IP来传递报文
- ❑ ICMP不是高层协议，而是IP层的协议

说明

- 类型：可看作是差错的种类，例如，3表示终点不可达；
- 代码：表明各种类型的不同具体情况，例如，类型为3时，代码
 - ▣ 0 = 网络不可达；
 - ▣ 1 = 主机不可达；
 - ▣ 2 = 协议不可用；
 - ▣ 3 = 端口不可达；
 - ▣
- 第5-8个字节，随类型的不同有所变化，例如，重定向消息中，用以表示网关的IP地址

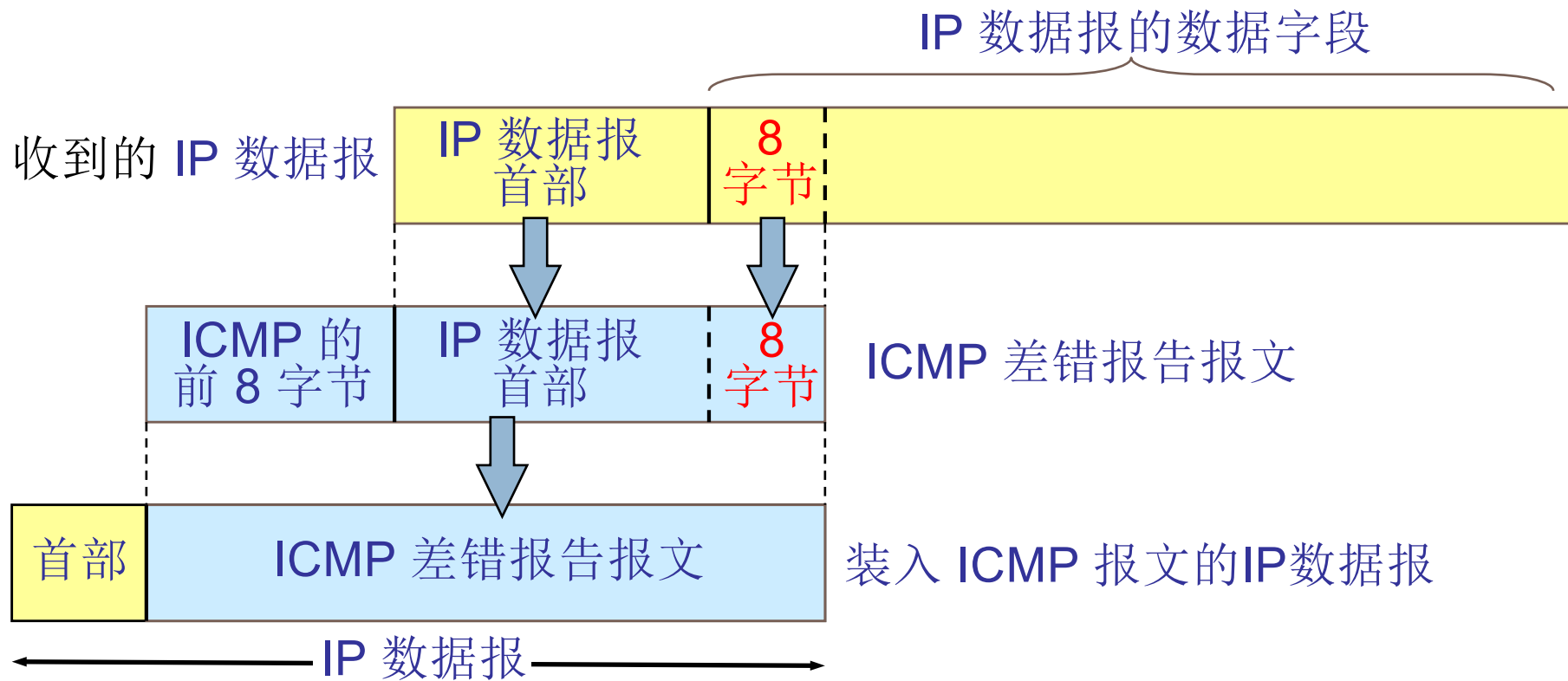
4.4.1 ICMP报文的种类

- ICMP 报文的种类有两种，即
 - ▣ ICMP 差错报告报文
 - ▣ ICMP 询问报文
- ICMP 报文
 - ▣ 4个字节是统一的格式，共有三个字段：即类型、代码和检验和
 - ▣ 接着的 4 个字节的内容与ICMP的类型有关

ICMP差错报告报文共有5种

- 终点不可达
- 源点抑制(Source quench)
 - ▣ 路由器或者主机发生拥塞，希望源点放慢发送速率时发送
- 时间超过
 - ▣ 路由器收到TTL=0的数据报
 - ▣ 终点在规定时间内收不到一个数据报的所有分片时发送
- 参数问题
 - ▣ 收到的数据报的字段的值错误时，丢弃数据报，向源点报告
- 改变路由（重定向）(Redirect)
 - ▣ 路由器告知主机改变的路由

ICMP差错报告报文的数据字段的内容



IP数据报的数据字段的前8个字节:运输层的端口号 (TCP和UDP) 以及运输层报文的发送序号 (TCP)

不应发送 ICMP 差错报告报文的几种情况

- 对 ICMP 差错报告报文不再发送 ICMP 差错报告报文。
- 对第一个分片的数据报片的所有后续数据报片都不发送 ICMP 差错报告报文
- 对具有多播地址的数据报都不发送 ICMP 差错报告报文
- 对具有特殊地址（如 127.0.0.0 或 0.0.0.0）的数据报不发送 ICMP 差错报告报文

ICMP询问报文

□ 回送请求/回送回答报文

- ▣ 主机或者路由器向目的主机发出回送请求报文，目的主机必须发回送回答报文，以测试可达性及其有关状态。

□ 时间戳请求或回答报文

- ▣ 请求或者回答当前的日期和时间，以进行时间同步或者测量时间。

4. 4. 2 ICMP的应用举例

PING

- 用来测试两个主机之间的连通性
- 使用了 ICMP 回送请求与回送回答报文
- 是应用层直接使用网络层ICMP的例子
 - ▣ 没有通过运输层的TCP或UDP

PING 的应用举例

```
C:\Documents and Settings\XXR>ping mail.sina.com.cn

Pinging mail.sina.com.cn [202.108.43.230] with 32 bytes of data:

Reply from 202.108.43.230: bytes=32 time=368ms TTL=242
Reply from 202.108.43.230: bytes=32 time=374ms TTL=242
Request timed out.
Reply from 202.108.43.230: bytes=32 time=374ms TTL=242

Ping statistics for 202.108.43.230:
    Packets: Sent = 4, Received = 3, Lost = 1 (25% loss),
    Approximate round trip times in milli-seconds:
        Minimum = 368ms, Maximum = 374ms, Average = 372ms
```

Traceroute 的应用举例

```
C:\Documents and Settings\XXR>tracert mail.sina.com.cn
```

```
Tracing route to mail.sina.com.cn [202.108.43.230]  
over a maximum of 30 hops:
```

1	24 ms	24 ms	23 ms	222.95.172.1
2	23 ms	24 ms	22 ms	221.231.204.129
3	23 ms	22 ms	23 ms	221.231.206.9
4	24 ms	23 ms	24 ms	202.97.27.37
5	22 ms	23 ms	24 ms	202.97.41.226
6	28 ms	28 ms	28 ms	202.97.35.25
7	50 ms	50 ms	51 ms	202.97.36.86
8	308 ms	311 ms	310 ms	219.158.32.1
9	307 ms	305 ms	305 ms	219.158.13.17
10	164 ms	164 ms	165 ms	202.96.12.154
11	322 ms	320 ms	2988 ms	61.135.148.50
12	321 ms	322 ms	320 ms	freemail43-230.sina.com [202.108.43.230]

```
Trace complete.
```

Traceroute原理

- Traceroute从源主机向目标主机发送一串IP数据报，封装的是无法交付的UDP数据报
- 数据报P1的TTL设为1
 - ▣ 路由器每接到数据报，对TTL要减1，若TTL为0，则丢弃，并发送ICMP时间超过差错报告报文；
- 源主机重发，发送数据报P_i，并设TTL为i
-
- 目标主机收到数据报后，不转发，也不把TTL减1，因为数据报无法交付，因此向源主机发送ICMP源点不可达差错报告报文

- 
- 36 若路由器R因为拥塞丢弃IP分组，则此时R可以向发出该IP分组的源主机发送的ICMP报文件类型是（ ）
 - ▣ A: 路由重定向
 - ▣ B: 目的不可达
 - ▣ C: 源抑制
 - ▣ D: 超时

4.5 因特网的路由选择协议

4.5.1 基本概念

- 静态路由配置 & 路由选择
- 理想的路由算法
 - ▣ 正确的和完整
 - ▣ 计算简单
 - ▣ 适应通信量和网络拓扑的变化，即自适应性
 - ▣ 稳定性、公平的、最佳的
- 有没有一种绝对的最佳路由算法？

不存在“最佳路由”？

- 不存在一种绝对最佳的路由选择算法
 - 所谓“最佳”，只能是相对于某一种特定要求的较为合理的选择，实际的路由选择算法，应尽可能接近于理想的算法
- 原因
 - ▣ 路由选择是个非常复杂的问题
 - ▣ 它是网络中的所有结点共同协调工作的结果
 - ▣ 路由选择的环境往往是不不断变化的，而这种变化有时无法事先知道。

从路由算法的自适应性考虑

□ 静态路由选择策略

- ▣ 即非自适应路由选择
- ▣ 其特点是简单和开销较小，但不能及时适应网络状态的变化。

□ 动态路由选择策略

- ▣ 即自适应路由选择
- ▣ 其特点是能较好地适应网络状态的变化，但实现起来较为复杂，开销也比较大。

因特网路由选择面对的问题

- 因特网的规模非常大，带来如下问题：
 - ▣ 如果让所有的路由器知道所有的网络应怎样到达，则这种路由表将非常大，处理起来也太花时间
 - ▣ 所有这些路由器之间交换路由信息所需的带宽就会使因特网的通信链路饱和。
 - ▣ 另外，许多单位不愿意外界了解其网络的布局细节和采用的路由选择协议，但希望连接到因特网上。
- 因此，因特网采用分层次的路由选择协议。

举例：分层次路由选择的好处

- 例如

- ▣ 100个结点的网络，分成10个区域，每个区域10个结点

- 一般路由算法

- ▣ 每个结点存放100条信息

- 层次路由

- ▣ 只存放本区域内10个结点的信息，以及另外9个区域的信息
 - ▣ 节约空间，减少交换信息

自治系统 AS (Autonomous System)

□ 自治系统的定义

- ▣ 在单一的技术管理下的一组路由器
- ▣ 这些路由器使用一种AS内部的路由选择协议和共同的度量以确定分组在该AS内的路由
- ▣ 同时，使用一种AS之间的路由选择协议用以确定分组在AS之间的路由。

□ AS强调以下事实

- ▣ 尽管一个AS使用了多种内部路由选择协议和度量，对其他AS表现出的是一个单一的和一致的路由选择策略。

因特网有两大类路由选择协议

- 内部网关协议IGP (Interior Gateway Protocol)
 - ▣ 即在一个自治系统内部使用的路由选择协议
 - ▣ 这类路由选择协议使用最多，如 RIP 和 OSPF 协议
- 外部网关协议EGP(External Gateway Protocol)
 - ▣ 若源站和目的站处在不同的自治系统中，当数据报传到一个自治系统的边界时，就需要使用一种协议将路由选择信息传递到另一个自治系统中；
 - ▣ 该类协议就是外部网关协议EGP。在外部网关协议中目前使用最多的是 BGP-4。

自治系统和 内部网关协议、外部网关协议



自治系统之间的路由选择也叫做**域间路由选择**(interdomain routing);
在自治系统内部的路由选择叫做**域内路由选择**(intradomain routing)

注意

- 路由器&网关

- ▣ 因特网的早期RFC文档中未使用“路由器”而是使用“网关”这一名词。

- IGP和EGP 是协议类别的名称

- ▣ 但RFC在使用EGP这个名词时出现了一点混乱
 - ▣ 最早的一个外部网关协议的协议名字正好也是EGP

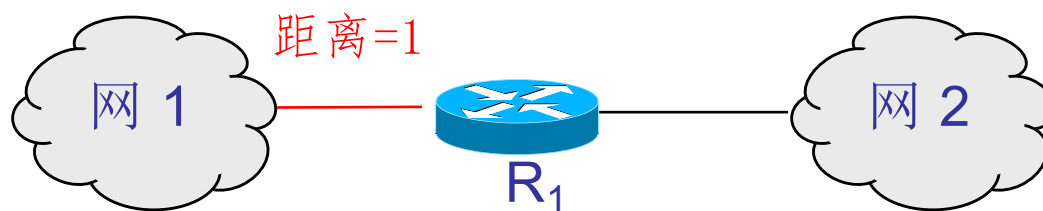
4.5.2 内部网关协议 RIP (Routing Information Protocol)

- 生活中的例子
 - ▣ 出行路线规划
- 基本思想
 - ▣ RIP是一种分布式的基于距离向量的路由选择协议
 - ▣ RIP协议要求网络中的每一个路由器都要维护从它自己到其他每一个目的网络的距离记录
- RIP是内部网关协议中最先得到广泛使用的协议

“1. 距离” 的定义

□ 距离

- 从一路由器到直接连接的网络的距离定义为 1
- 从一个路由器到非直接连接的网络的距离定义为所经过的路由器数加 1。



□ 距离的含义

- RIP 协议中的“距离”也称为“跳数” (hop count)，因为每经过一个路由器，跳数就加1
- 这里的“距离”实际上指的是“最短距离”

RIP的路由评价标准

- RIP认为一个好的路由：
 - ▣ 通过的路由器的数目少，即“距离短”
- RIP允许一条路径最多只能包含15个路由器
 - ▣ “距离”为16时即相当于不可达
 - ▣ 因此，RIP只适用于小型互联网
- RIP能否在两个网络之间同时使用多条路由？
 - ▣ 不能
 - ▣ RIP选择一个具有最少路由器的路由，哪怕还存在另一条高速但路由器较多的路由。

路由表的建立

□ 建立思想

- ▣ 路由器开始工作时，只知道到直接连接的网络的距离（距离=1）
- ▣ 以后，每一个路由器也只和相邻路由器交换并更新路由信息
- ▣ 最终，所有路由器都会知道到达本AS中任何一个网络的最短距离和下一跳路由器的地址

□ 优点

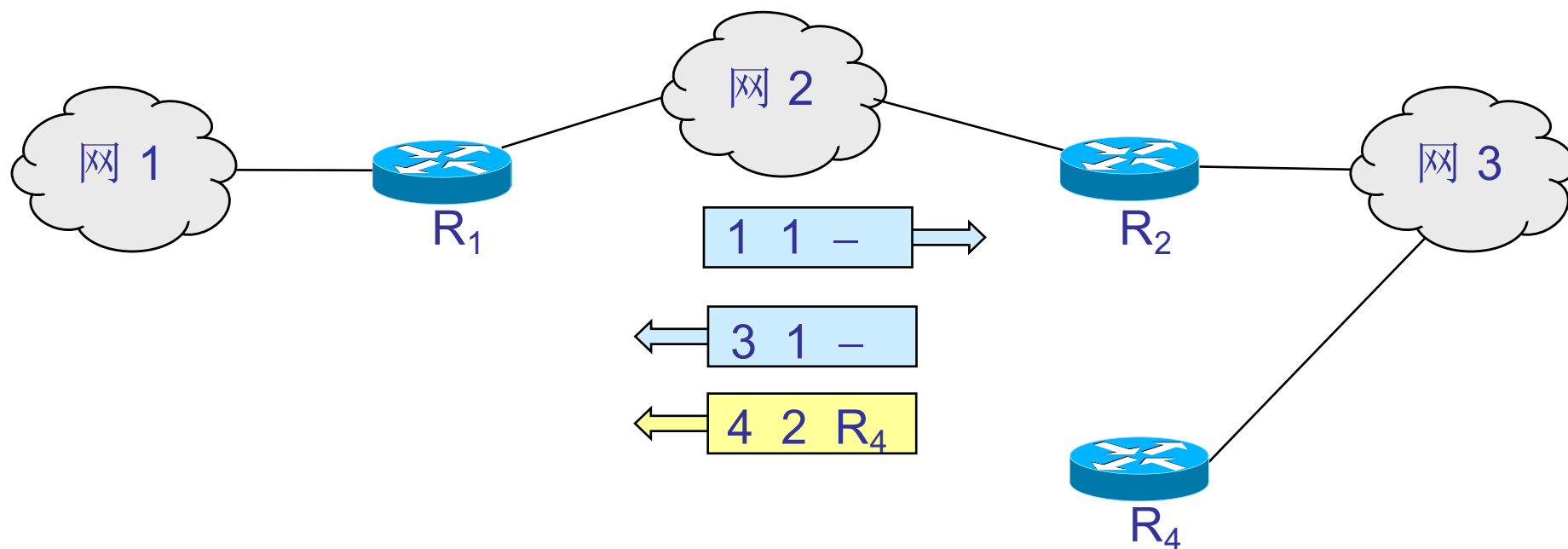
- ▣ RIP 协议的**收敛**(convergence)过程较快
- ▣ 即，在AS中所有的结点都得到正确的路由选择信息的过程

2. 核心：距离向量算法

□ 路由项的构成

□ 例如：Net₁ 3 X

□ 表示当前路由器，通过下一跳（路由器X），到达Net₁的距离是3



距离向量算法

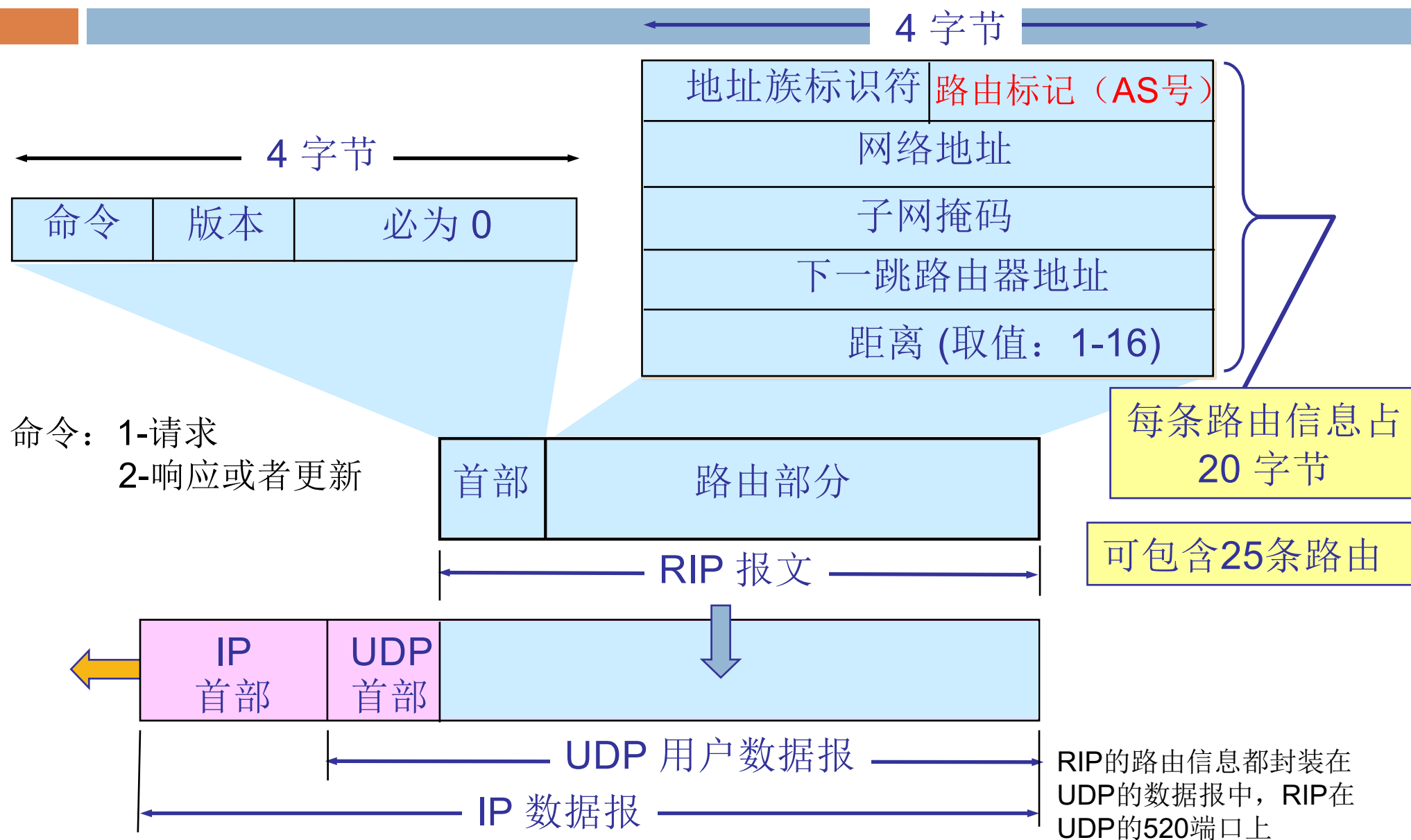
- 1. 收到相邻路由器X的一个RIP报文
- 2. 预处理
 - ▣ 修改此RIP报文中的所有项目, “下一跳” = X, “距离” + 1
- 3. 对修改后的RIP报文中的每一个项目:
 - ▣ 若 项目中的目的网络不在路由表中 **新路由项**
 - ▣ 则 把该项目加到路由表中
 - ▣ 否则
 - 若 下一跳字段给出的路由器地址是同样的 **路由改变**
 - 则 用收到的项目替换原路由表中的项目;
 - 否则
 - 若 收到项目中的距离小于路由表中的距离 **更短路由**
 - 则 进行更新;
- 4. 若3分钟未收到相邻路由器的更新报文, 此路由器不可达, 距离=16
- 5. 返回

RIP协议的三个要点

- 一个路由器仅和相邻路由器交换信息
- 交换的信息是当前本路由器所知道的全部信息，
即自己的路由表
- 按固定的时间间隔交换路由信息，例如，每隔 30 秒

使得每一个路由器到每一个目的网络的路由最短

3. RIP2 协议的报文格式



RIP2报文的路由部分

□ 构成

- ▣ 若干个路由信息，每个路由信息需要用20个字节
- ▣ 地址族标识符（又称为地址类别）字段用来标志所使用的地址协议

□ 路由标记填入自治系统的号码

- ▣ 这是考虑RIP有可能收到本自治系统以外的路由选择信息

RIP协议的优缺点

□ 优点

- ▣ 实现简单、实现开销较小、收敛快

□ 存在的问题

- ▣ 当网络出现故障时，要经过比较长的时间才能将此信息传送到所有的路由器；
- ▣ RIP限制了网络的规模，最大距离为15（16 表示不可达）；
- ▣ 路由器之间交换的路由信息是路由器中的完整路由表，因而随着网络规模的扩大，开销也就增加。

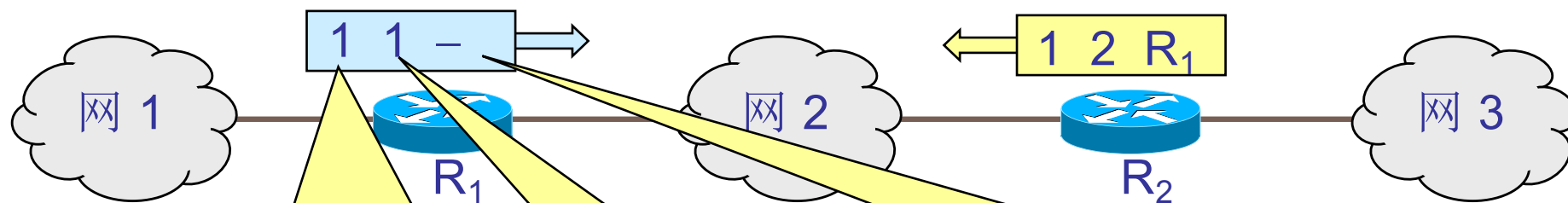
实际应用中的缺陷

□ 计数到无穷问题

- ▣ 对好消息能迅速做出反应，而对坏消息的反馈则非常缓慢
- ▣ 在一次向量交换中，好消息就被处理了
- ▣ 例如：设想由一个路由器到目标x的最短距离很大，假定该路由器在下一次与邻居交换向量表时获知，由邻居A到达x有一更短的延迟，则该路由器将选择通过A发送信息到x的路径

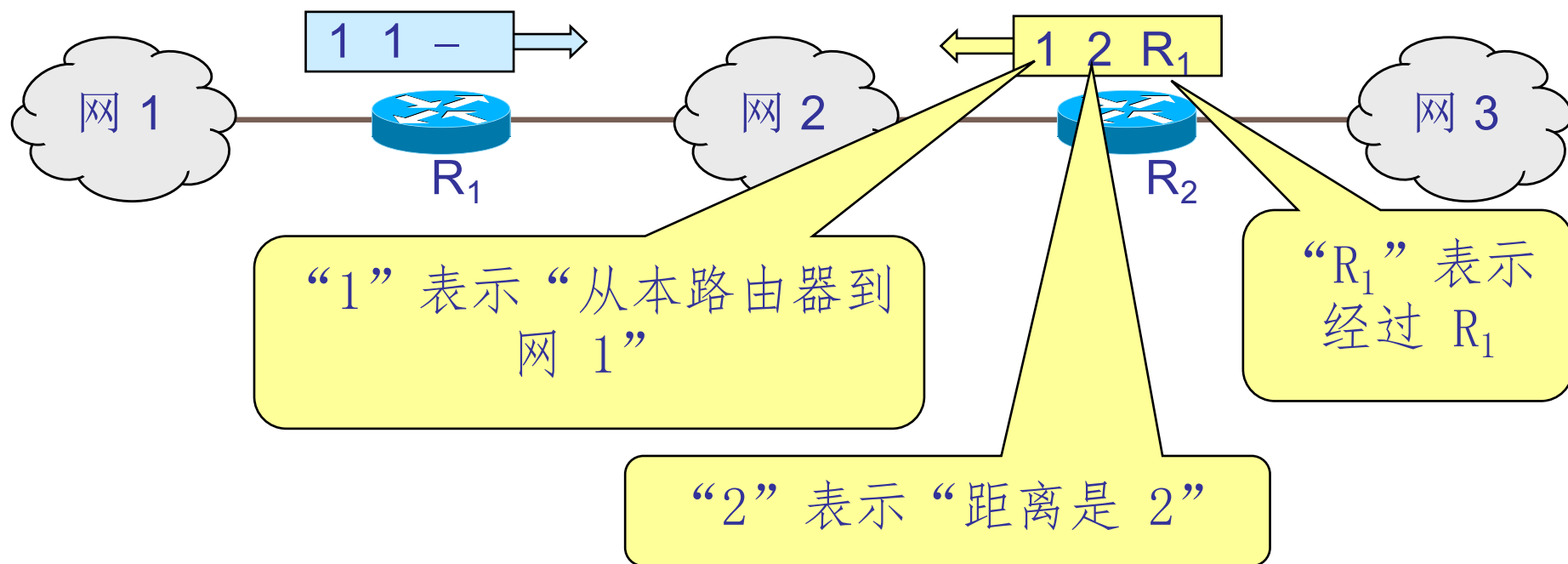
□ 另外一种情况：

正常情况



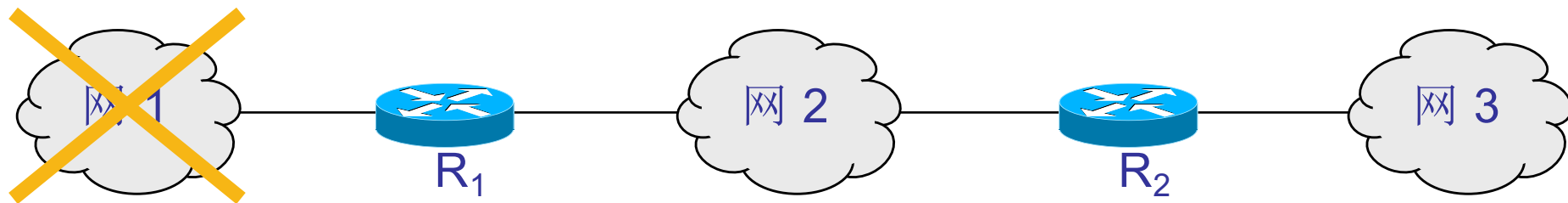
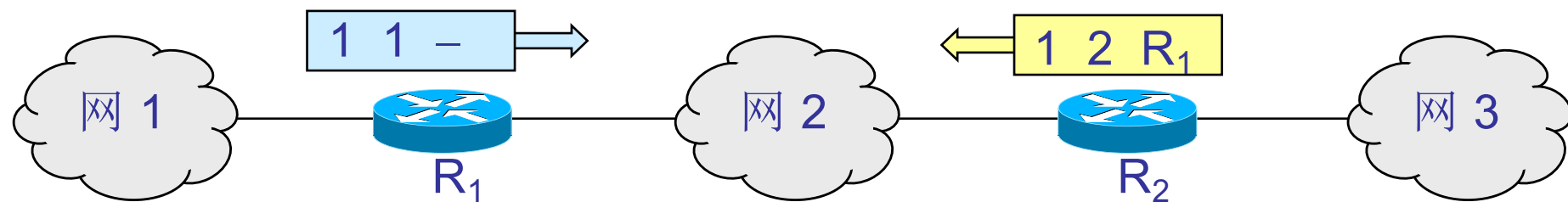
R₁ 说：“我到网 1 的距离是 1，是直接交付。”

正常情况



R_2 说：“我到网 1 的距离是 2，是经过 R_1 。”

正常情况



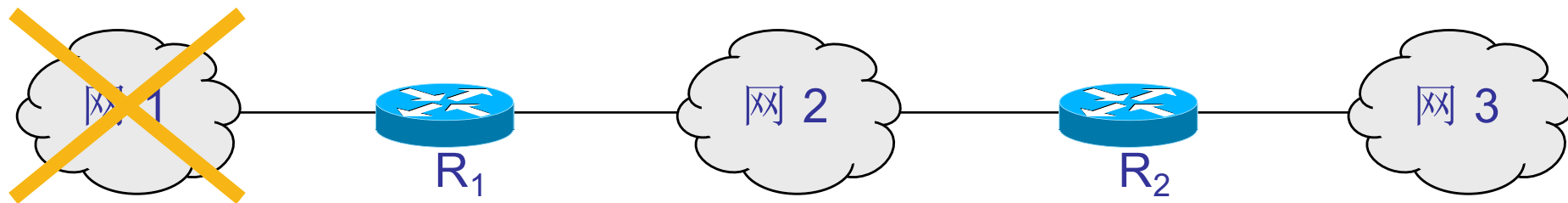
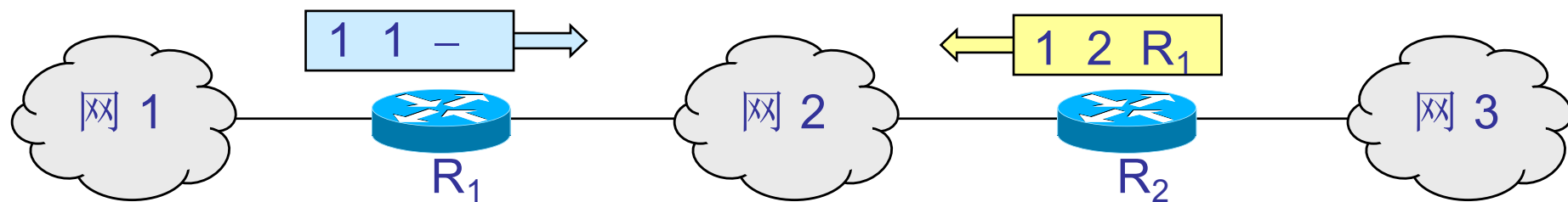
网 1 出了故障



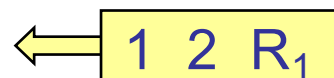
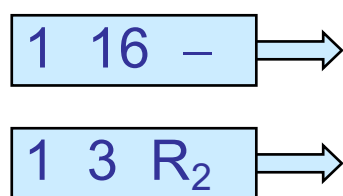
R₁ 说：“我到网 1 的距离是 16 （表示无法到达），是直接交付。”

但 R₂ 在收到 R₁ 的更新报文之前，还发送原来的报文，因为这时 R₂ 并不知道 R₁ 出了故障。

正常情况

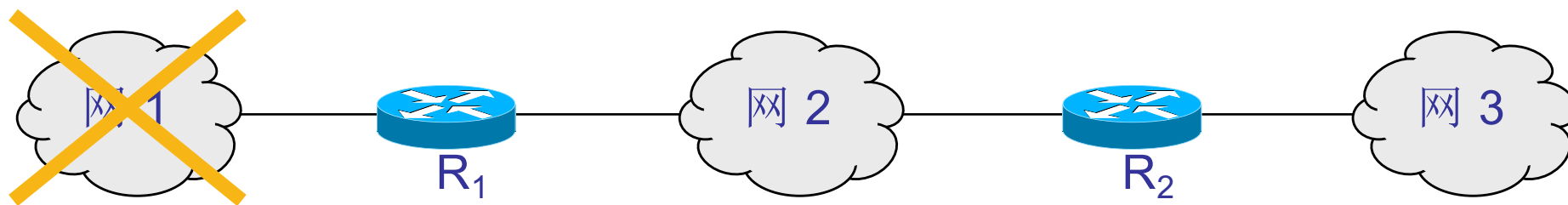
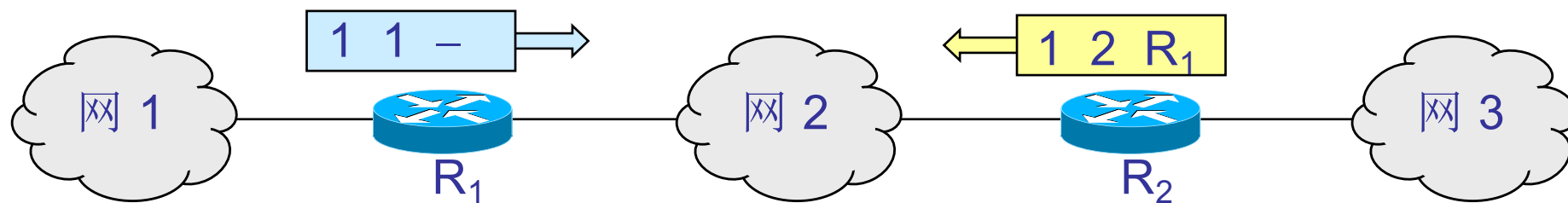


网 1 出了故障



R_1 收到 R_2 的更新报文后，误认为可经过 R_2 到达网 1，于是更新自己的路由表，说：“我到网 1 的距离是 3，下一跳经过 R_2 ”。然后将此更新信息发送给 R_2 。

正常情况



网 1 出了故障



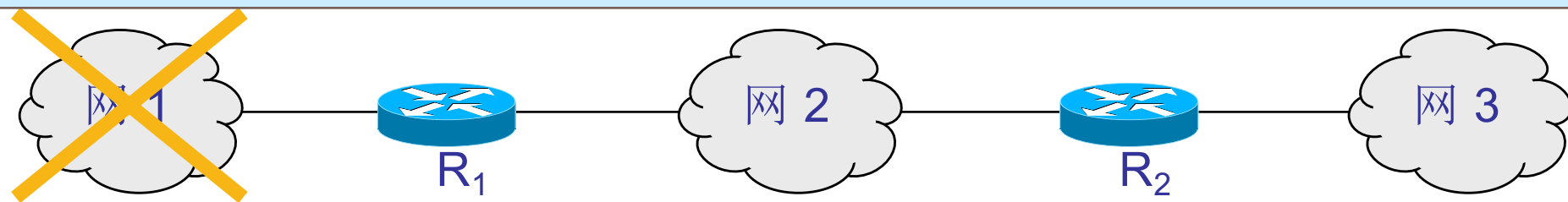
R₂ 以后又更新自己的路由表为 “1, 4, R₁”，表明 “我到网 1 距离是 4，下一跳经过 R₁”。

正
常

1 1 - →

← 1 2 R₁

这就是好消息传播得快，而坏消息传播得慢。
网络出故障的传播时间往往需要较长的时间(例如数分钟)。
这是 RIP 的一个主要缺点。



网1出了故障

1 16 - →

← 1 2 R₁

1 3 R₂ →

← 1 4 R₁

1 5 R₂ →

⋮

⋮

1 16 R₂ →

← 1 16 R₁

这样不断更新下去，直到 R₁ 和 R₂ 到网 1 的距离都增大到 16 时，
R₁ 和 R₂ 才知道网 1 是不可达的。

练习

- 35（2010）、某自治系统采用RIP协议，若该自治系统内的路由器R1收到其邻居路由器R2的距离矢量中包含信息<net1, 16>，则可能得出的结论是（ ）
 - ▣ A： R2可以经过R1到达net1，跳数为17
 - ▣ B： R2可以到达net1，跳数为16
 - ▣ C： R1可以经过R2到达net1，跳数为17
 - ▣ D： R1不能经过R2到达net1

习题：在某个网络中，R1和R2为相邻路由器，表1为R1的原路由表，表2为R2广播的距离向量报文，请应用RIP协议更新R1路由表，给出更新后的路由表。

表1. R1的原路由表

目的网络	距离	下一跳
10.0.0.0	0	直接交付
30.0.0.0	7	R7
40.0.0.0	3	R2
45.0.0.0	4	R8
180.0.0.0	5	R2
190.0.0.0	10	R5

表2. R2广播的距离向量报文

目的网络	距离
10.0.0.0	4
30.0.0.0	4
40.0.0.0	2
41.0.0.0	3
180.0.0.0	5

目的网络	距离	下一条



目的网络	距离	下一条
10.0.0.0	0	直接交付
30.0.0.0	5	R2
40.0.0.0	3	R2
41.0.0.0	4	R2
45.0.0.0	4	R8
180.0.0.0	6	R2
190.0.0.0	10	R5

Thanks, Any Question?

The End.

chenwh@nwnu.edu.cn

