



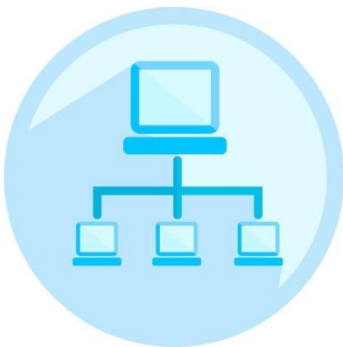
计算机网络



顾 军

计算机学院

jgu@cumt.edu.cn





专题4：数据包怎么在互联网中寻路和转发？



- 应用层(application layer)
- 运输层(transport layer)
- 网络层(network layer)
- 数据链路层(data link layer)
- 物理层(physical layer)





Q26: 分层次的因特网路由选择策略?

- 因特网的规模非常大。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。而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。
- 许多单位不愿意外界了解自己单位网络的布局细节和本部门所采用的路由选择协议（这属于本部门内部的事情），但同时还希望连接到因特网上。
- 因特网采用分层次的路由选择协议。





自治系统 AS

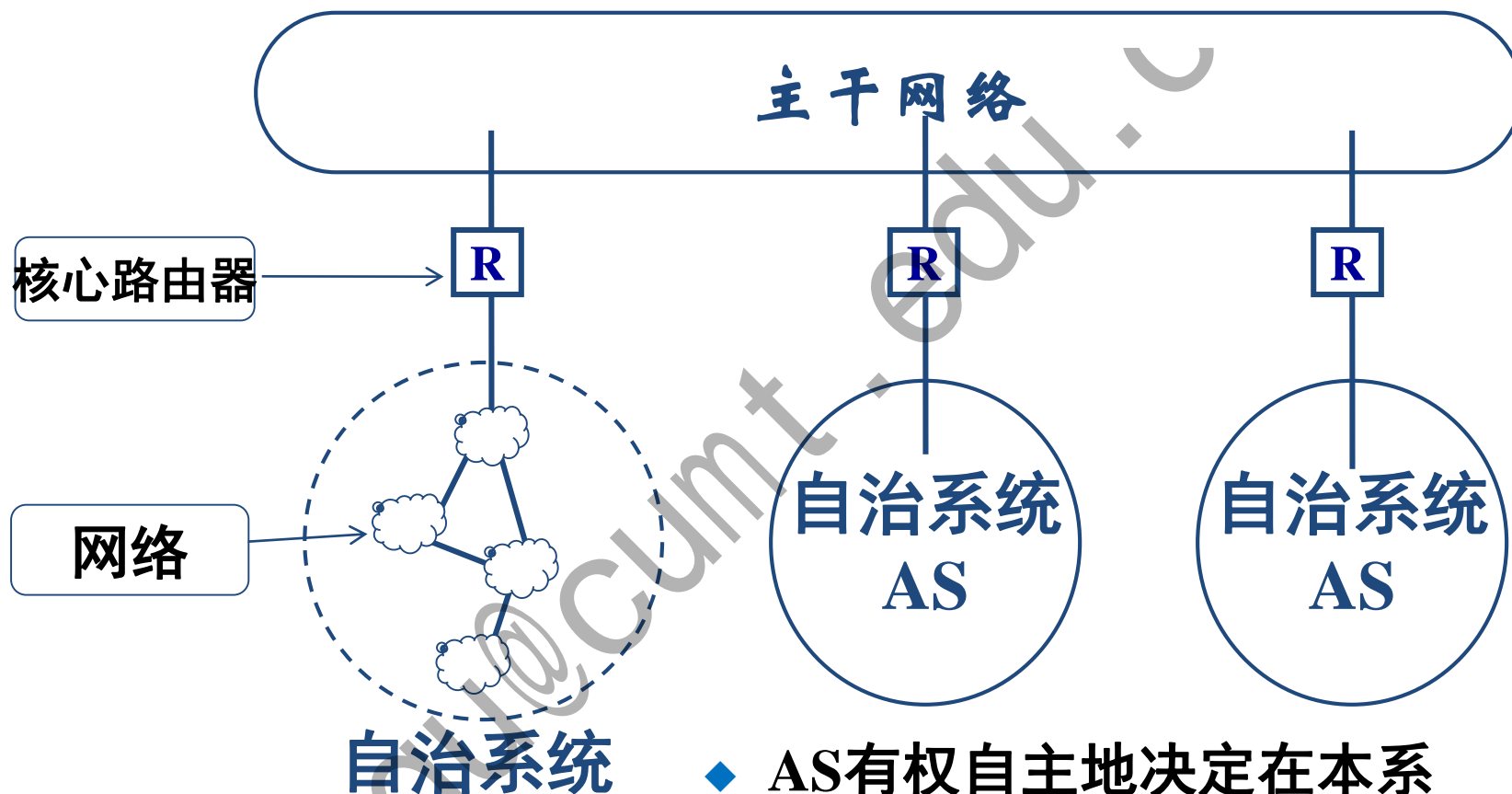
(Autonomous System)

- 自治系统 AS 的定义：在**单一的技术管理下**的一组路由器，而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，同时还使用一种 AS 之间的路由选择协议用以确定分组在 AS 之间的路由。
- 现在对自治系统 AS 的定义是强调下面的事实：尽管一个 AS 使用了多种内部路由选择协议和度量，但重要的是一个 AS 对其他 AS 表现出的是一个**单一的**和**一致的路由选择策略**。
- 一个 AS 的所有网络都由一个部门管辖（ISP、公司、大学，政府部门等）





Internet是由大量的AS组成，各个AS由不同机构操纵。

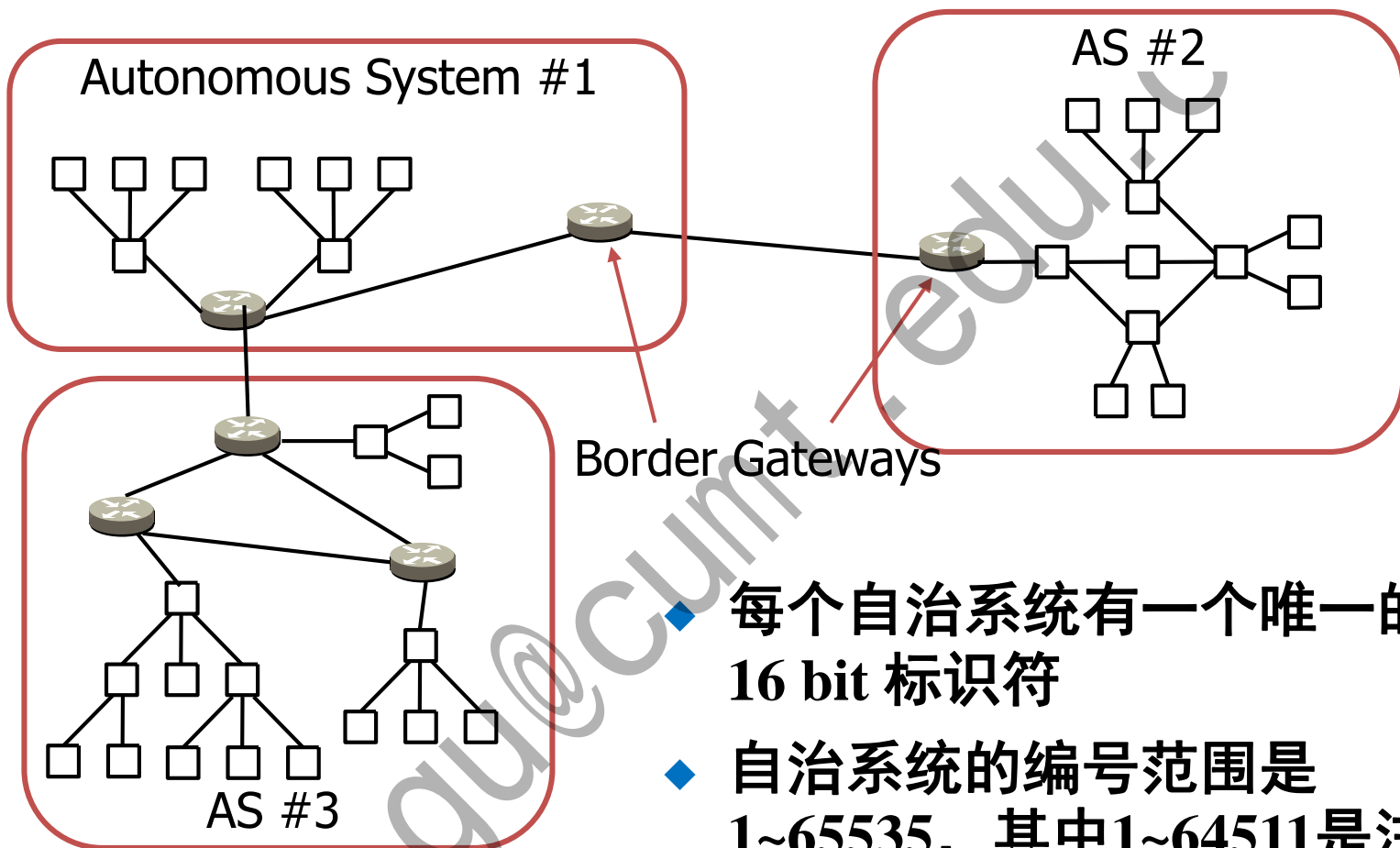


- ◆ AS有权自主地决定在本系统内采用何种路由协议
- ◆ 一个AS所有路由器在本AS内都必须是连通的





自治系统标识



- ◆ 每个自治系统有一个唯一的 16 bit 标识符
- ◆ 自治系统的编号范围是 1~65535，其中 1~64511 是注册的因特网编号，64512~65535 是专用网络编号





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)





Q27: 基于距离向量的RIP协议 ?

- RIP (Routing Information Protocol) 是一种分布式的基于距离向量的路由选择协议，是内部网关协议 IGP中最先得到广泛使用的协议
- 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。





“距离”的定义

- 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。
 - 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。
- RIP 认为一个好的路由就是它通过的路由器的数目少，即“距离短”，不一定是物理距离最短、或速率最高、或最快。
- RIP 允许一条路径最多只能包含 15 个路由器，“距离”的值为 16 时即相当于不可达。因此，RIP 只适用于小型互联网。





RIP 协议的三个要点

- 与谁交换：
 - 仅和相邻路由器交换信息。
- 交换什么：
 - 交换的信息是本路由器当前所知道的全部信息，即自己的路由表，即距离向量。
- 何时交换：
 - 定时交换路由信息，例如，每隔 30 秒。





RIP路由表的更新

更新

R1 路由表

Destin	Cost	Next-hop
N1	1	Direct

距离向量
更新消息

Destin	Cost
N1	1

R2 路由表

Destin	Cost	Next-hop
N1	2	R1

距离向量
更新消息

Destin	Cost
N1	2

N1

R1

R2

Destin Cost

N1

2

距离
向量
调整





距离向量算法

收到相邻路由器（其地址为 X）的一个 RIP 报文：

(1) 先修改此 RIP 报文中的所有项目：把“下一跳”字段中的地址都改为 X，并把所有的“距离”字段的值加 1。

(2) 对修改后的 RIP 报文中的每一个项目，重复以下步骤：
若项目中的目的网络 N 不在路由表中，则把该项目加到路由表中。
否则（即在路由表中有目的网络 N）。

 若下一跳字段给出的路由器地址是同样的 X，则把收到的项目替换原路由表中的项目。

 否则（即这个项目是到目的网络 N，但是下一跳不是 X）

 若 收到项目中的距离小于路由表中的距离，则进行更新，
 否则，什么也不做。

(3) 若 3 分钟还没有收到相邻路由器的更新路由表，则把此相邻路由器记为不可达路由器，即将距离置为 16（距离为 16 表示不可达）。

(4) 返回。



RIP路由表更新举例

【例】现在B收到其相邻路由器C发来的路由信息，请画出B更新后的路由表。

C的路由信息	
目的网络	距离
N ₂	3
N ₃	4
N ₆	5
N ₈	4
N ₉	5



B的路由表		
目的网络	距离	下一跳
N ₁	1	-
N ₂	2	C
N ₆	8	F
N ₈	4	E
N ₉	4	F

如同路由器一样，我们不需要知道该网络的拓扑。





现在B收到其相邻路由器C发来的路由信息，请画出B更新后的路由表。

修改C所有项目		
目的网络	距离	下一跳
N ₂	3+1=4	C
N ₃	4+1=5	C
N ₆	5+1=6	C
N ₈	4+1=5	C
N ₉	5+1=6	C

B的路由表		
目的网络	距离	下一跳
N ₁	1	-
N ₂	2	C
N ₆	8	F
N ₈	4	E
N ₉	4	F

B的路由表		
目的网络	距离	下一跳
N ₁	1	-
N ₂	3+1=4	C
N ₃	4+1=5	C
N ₆	5+1=6	C
N ₈	4	E
N ₉	4	F





RIP协议的工作过程

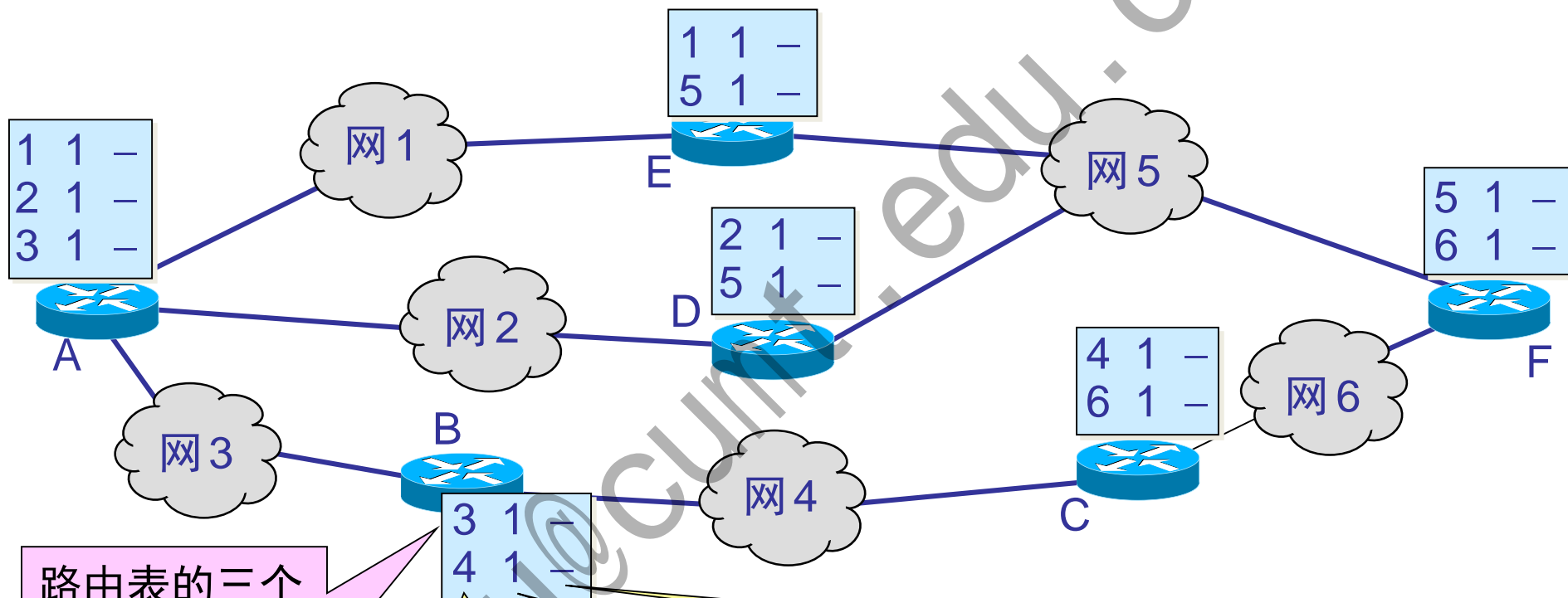
- 路由器在刚刚开始工作时，只知道到直接连接的网络的距离（此距离定义为1）。
- 以后，每一个路由器也只和数目非常有限的相邻路由器交换并更新路由信息。
- 经过若干次更新后，所有的路由器最终都会知道到达本自治系统中任何一个网络的最短距离和下一跳路由器的地址。
- RIP 协议的收敛(convergence)过程较快，即在自治系统中所有的结点都得到正确的路由选择信息的过程。





路由表建立过程

网络开始工作时，路由器只知道直接连接的网络，其距离为1，各路由表只有到相邻路由器的信息



路由表的三个主要字段：
目的网络，
跳数，
下一跳路由

“4”表示 “到网 4”的路由

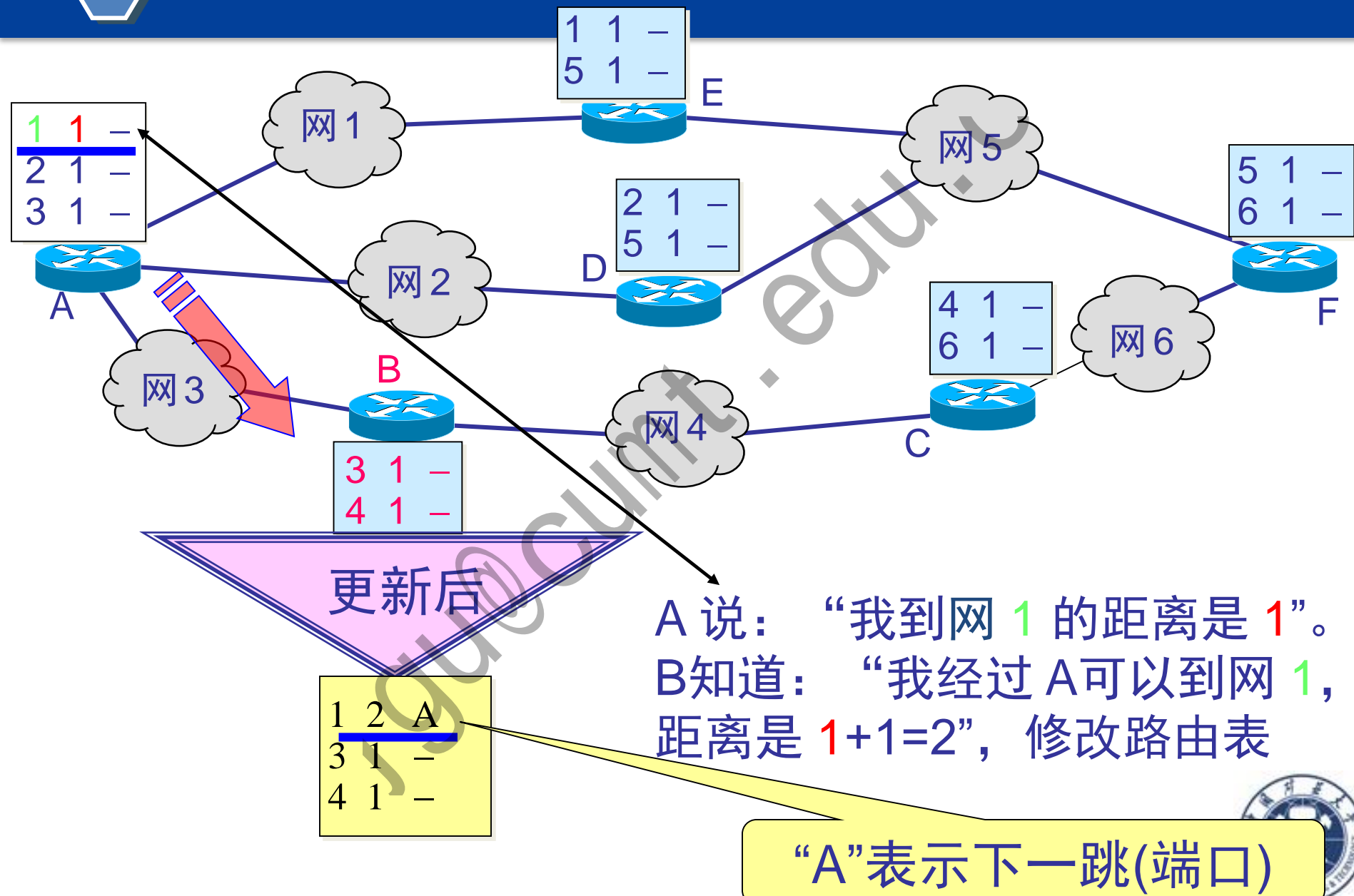
“-”表示 “直接交付”

“1”表示 “距离是 1”



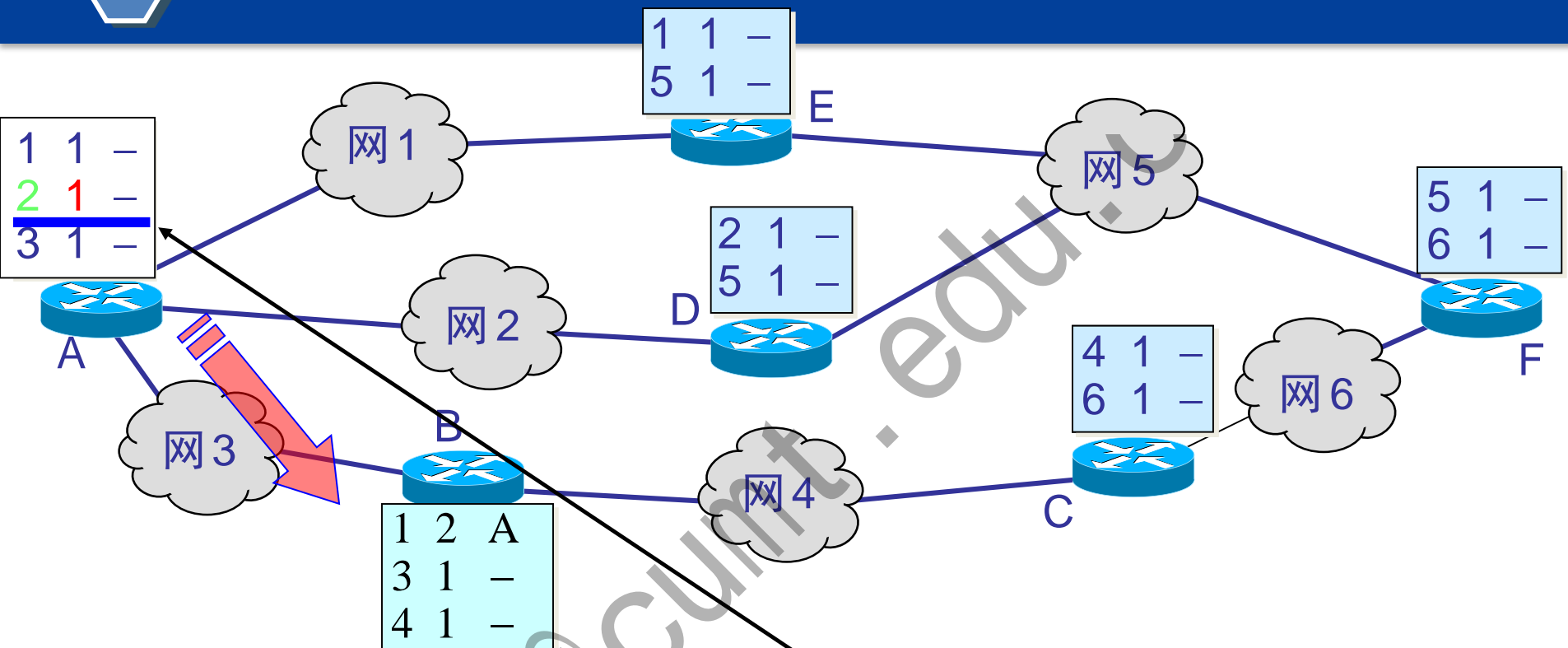


路由器 B 收到相邻路由器 A 的路由表—1





路由器 B 收到相邻路由器 A 的路由表-2



更新后

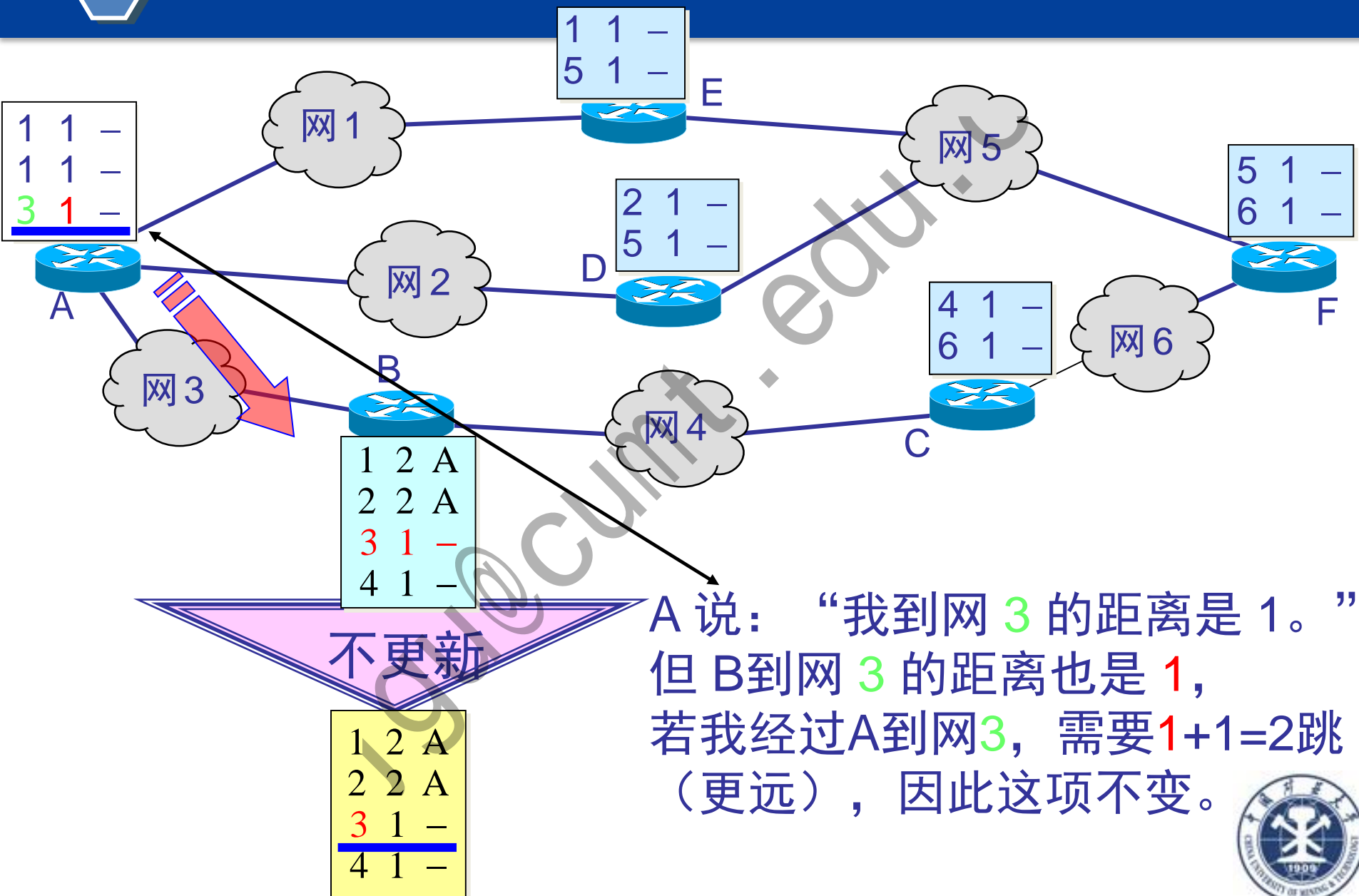
1	2	A
2	2	A
3	1	-
4	1	-

A 说：“我到网 2 的距离是 1。”
B 知道：我经过 A 可以到网 2，
距离是 $1+1=2$ ，修改路由表



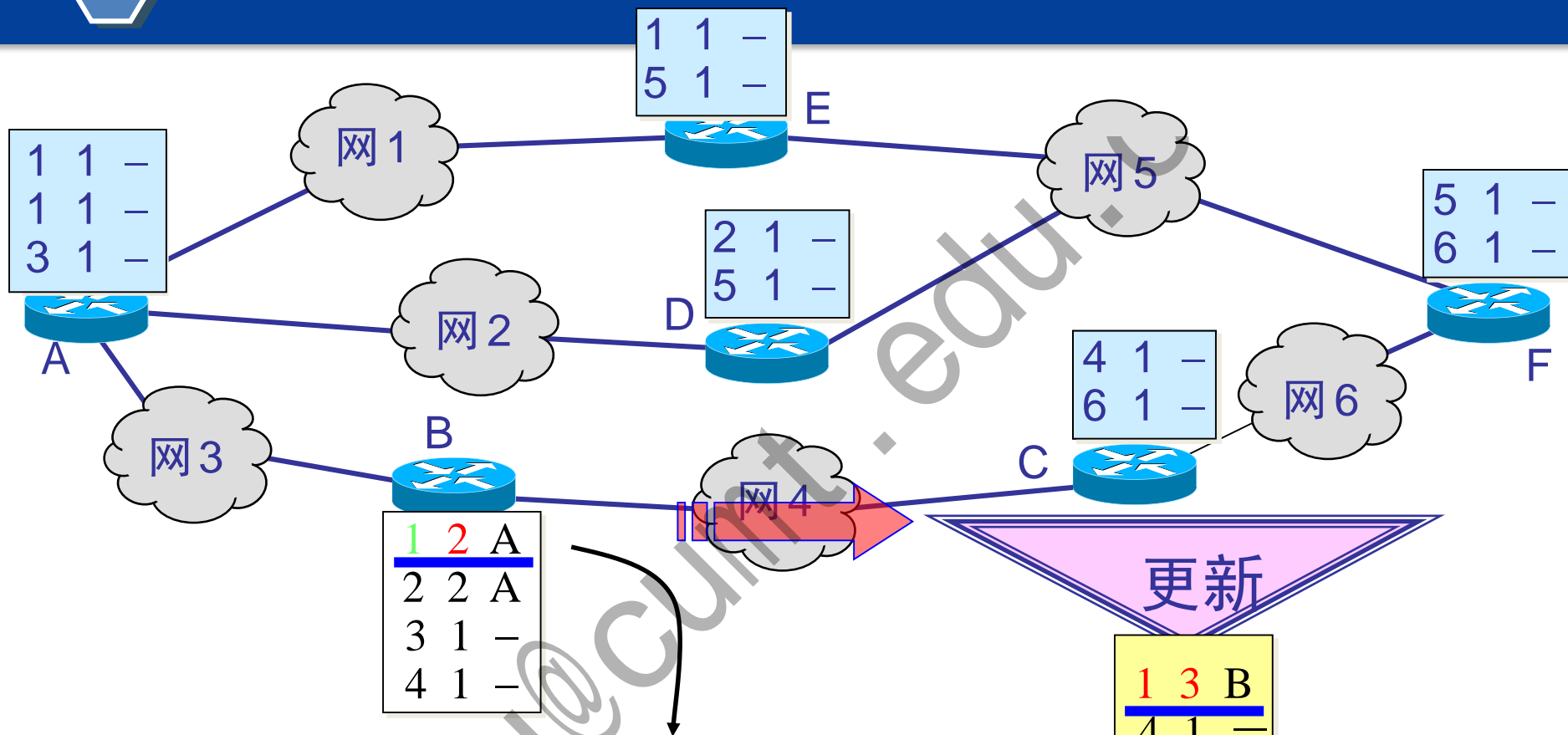


路由器 B 收到相邻路由器 A 的路由表—3





路由器 C 收到相邻路由器 AB 的路由表—1

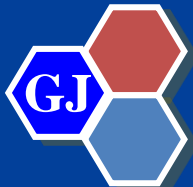


B说：“我到网 1 的距离是 2。”

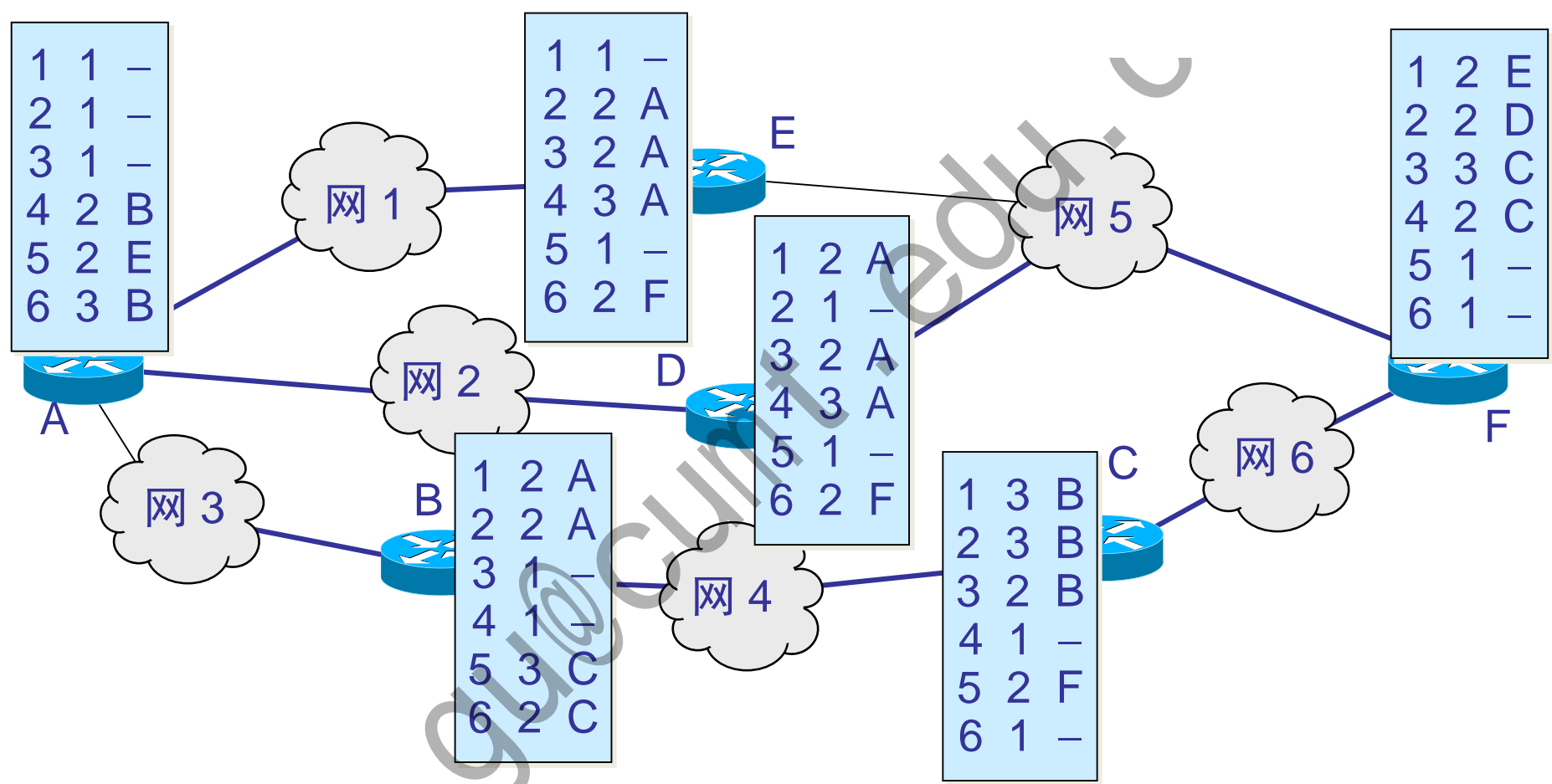
C知道：我经过 B 可以到网 1，

距离是 $2+1=3$ ：修改路由表，添加一条路由





经过若干次更新，所有路由器都知道到达本AS中任何一个网络的最短距离和下一跳路由器的地址。



1. 每一个路由器的位置不同，它们的路由表是不同的。
2. RIP 协议的收敛过程较快。



路由器之间交换信息

- **RIP**协议让互联网中的所有路由器都和自己的相邻路由器不断交换路由信息，并不断更新其路由表，使得从每一个路由器到每一个目的网络的路由都是最短的（即跳数最少）。
- 虽然所有的路由器最终都拥有了整个自治系统的全局路由信息，但由于每一个路由器的位置不同，它们的路由表也是不同的。
- **RIP** 不能在两个网络之间同时使用多条路由。**RIP** 选择一个具有最少路由器的路由（即最短路由），哪怕还存在另一条高速(低时延)但路由器较多的路由。





RIP 报文的传输

- RIP 协议使用传输层的UDP进行传送（使用 UDP 的端口 520）。
 - 形式上，RIP 协议的位置应当在应用层。
 - 逻辑上，是IP协议的配套，在网络层。



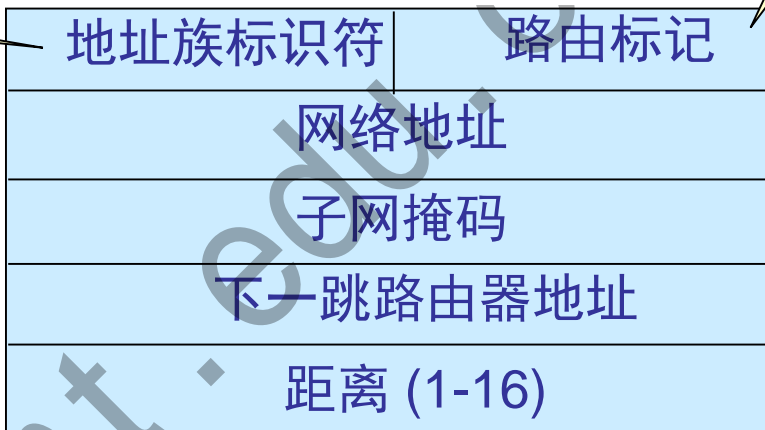


RIP2 协议的报文格式

AS号

又称为地址类别
IP地址则值为2

4 字节



4 字节

命令	版本	必为 0
----	----	------

“1”=请求路由信息
“2”=响应路由信息

首部	路由部分
----	------

RIP 报文

路由信息
(20 字节/路由)
可重复出现
最多 25 个

IP
首部

UDP
首部

UDP 用户数据报

IP 数据报





RIP2 报文

- 一个 RIP 报文最多可包括 25 个路由，因而 RIP 报文的最大长度是 $4 + 20 \times 25 = 504$ 字节。如超过，必须再用一个 RIP 报文来传送。
- **RIP2 具有简单的鉴别功能。**
 - 若使用鉴别功能，则将原来写入第一个路由信息（20字节）的位置用作鉴别。
 - 在鉴别数据之后才写入路由信息，但这时最多只能再放入 24 个路由信息。





RIP 协议的优缺点

- 优点:

- 实现简单，开销较小。

- 缺点:

- **RIP** 限制了网络的规模，它能使用的最大距离为 **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。
 - “坏消息传播得慢”，使更新过程的收敛时间过长。





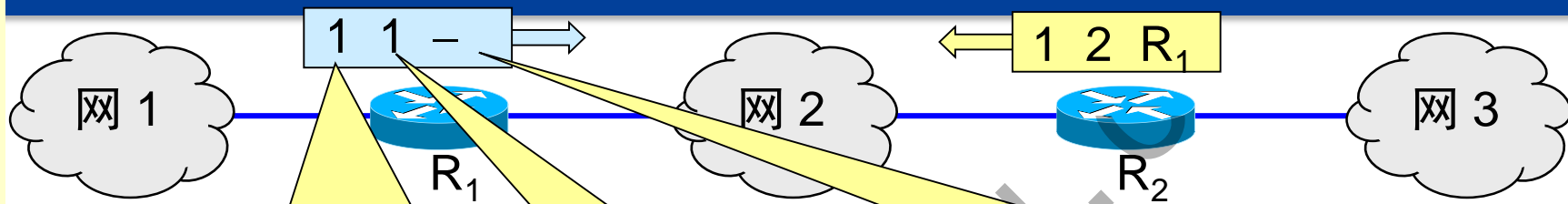
好消息传播得快，坏消息传播得慢

- **RIP协议特点：**好消息传播得快，坏消息传播得慢。
- **RIP存在的一个问题：**当网络出现故障时，要经过比较长的时间(例如数分钟)才能将此信息传送到所有的路由器。





正常情况



“1”表示“从本路由器到网 1”

“-”表示“直接交付”

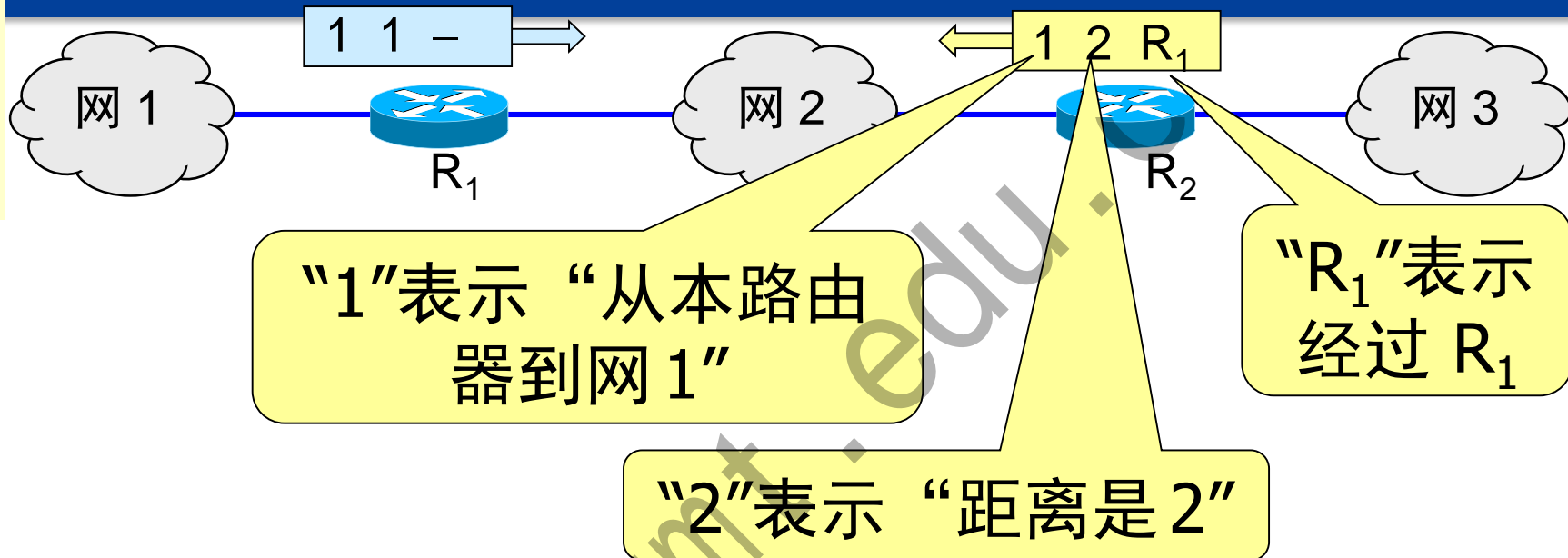
“1”表示“距离是 1”

R_1 说：“我到网 1 的距离是 1，是直接交付。”





正常情况

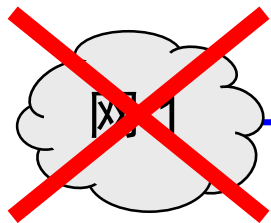
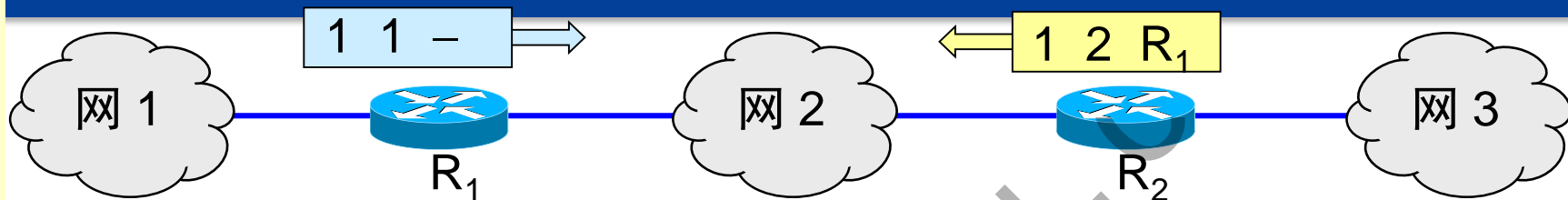


R_2 说：“我到网 1 的距离是 2，是经过 R_1 。”





正常情况



网 1 出了故障



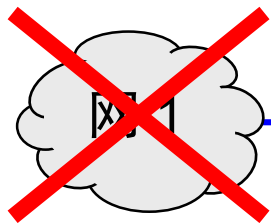
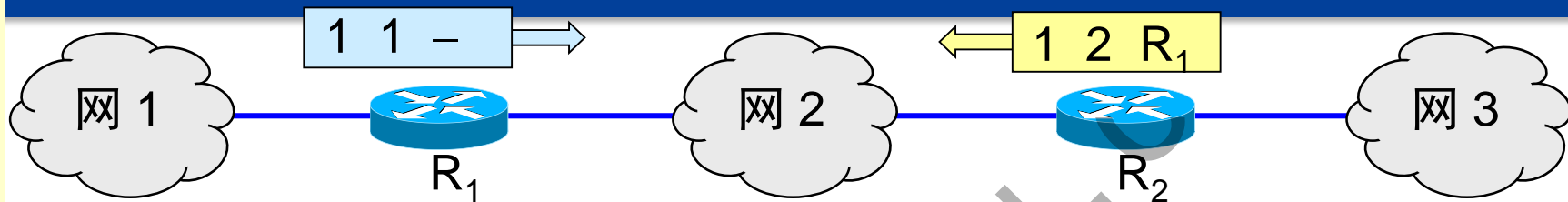
R_1 说：“我到网 1 的距离是 16（表示无法到达），是直接交付。”

但 R_2 在收到 R_1 的更新报文之前，还发送原来的报文，因为这时 R_2 并不知道 R_1 出了故障。

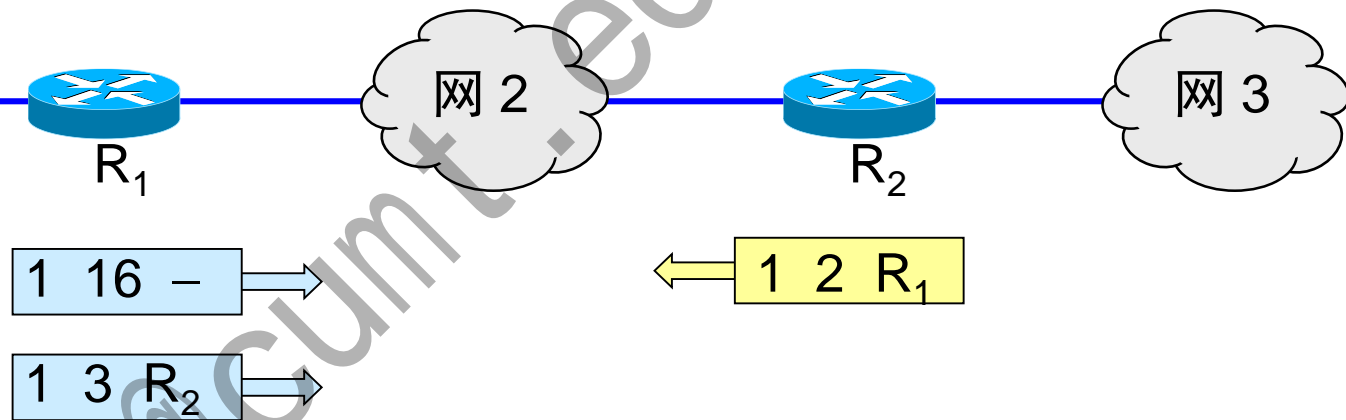




正常情况



网 1 出了故障

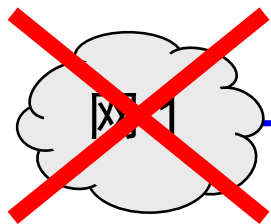
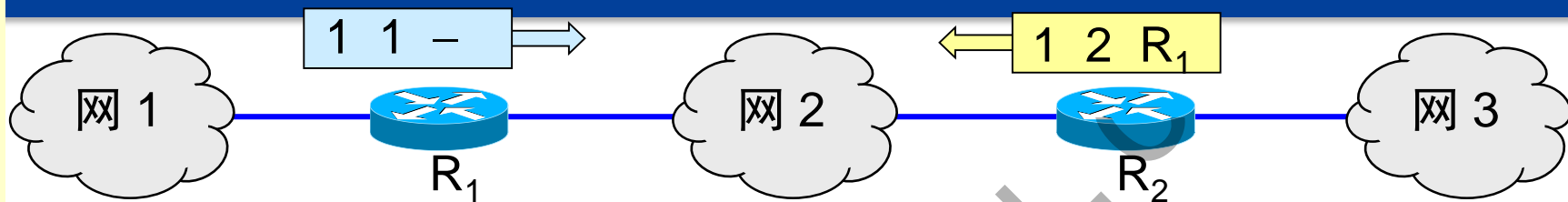


R_1 收到 R_2 的更新报文后，误认为可经过 R_2 到达网 1，于是更新自己的路由表，说：“我到网 1 的距离是 3，下一跳经过 R_2 ”。然后将此更新信息发送给 R_2 。





正常情况



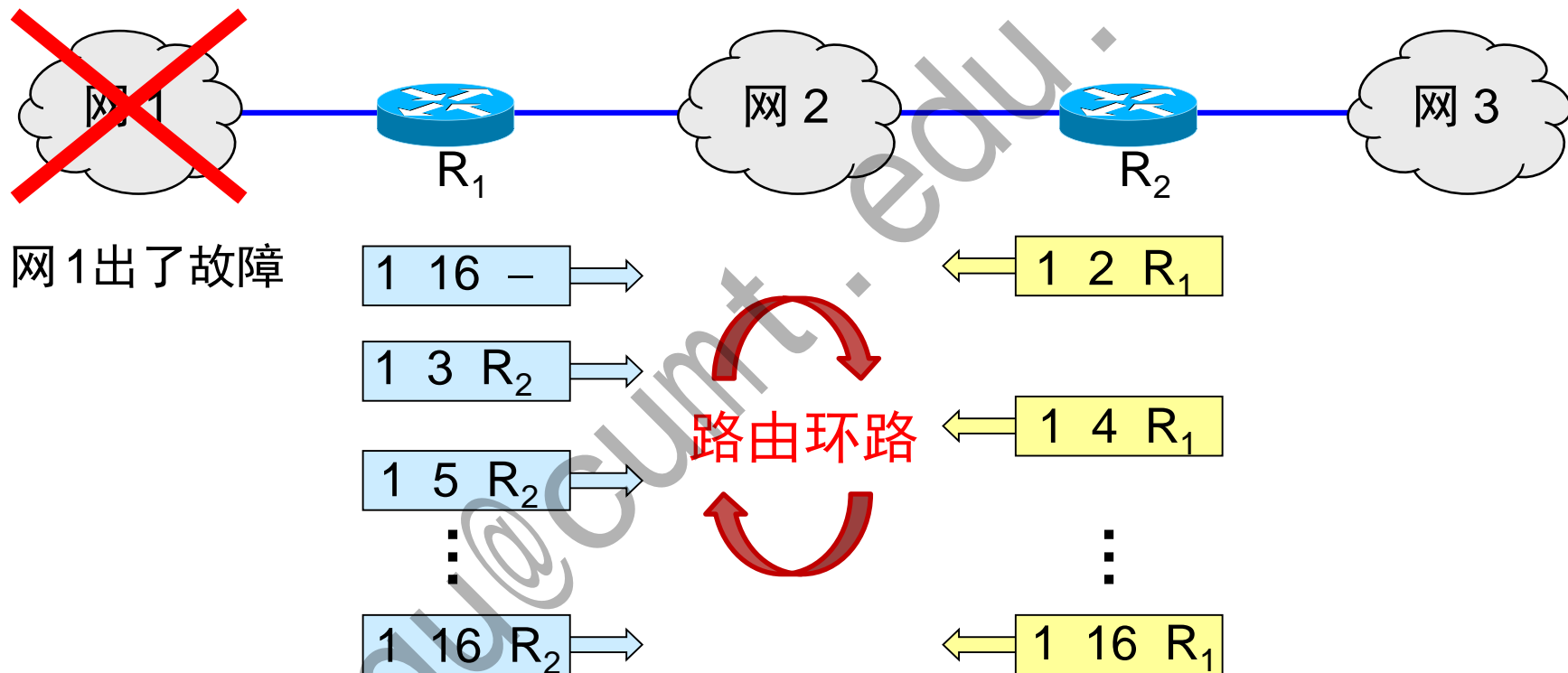
网 1 出了故障



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”。



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。 又称“**流言蜚语路由**”。



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。



Q28: RIP 的防环机制?

- 在维护路由表信息的时候，如果在拓扑发生改变后，网络收敛缓慢产生了不协调或者矛盾的路由选择条目，就会发生**路由环路**问题。
- 在这种条件下，路由器对无法到达的网络路由不予理睬，导致用户的数据包不停在网络上循环发送，最终造成网络资源的严重浪费。

忠言逆耳，选择性失聪



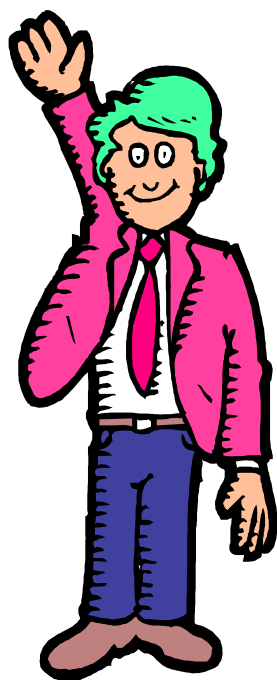


RIP 的防环机制

六种同时使用才能完全
解决路由环路

- 1-最大度量值（maximum hop count）：定义最大跳数（最大为15跳），当跳数为16跳时，目标为不可达。
- 2-水平分割（split horizon）：从一个接口学习到的路由不会再广播回该接口。
- 3-路由毒化（route poisoning）：当拓扑变化时，路由器会将失效的路由标记为possibly down状态，并分配一个不可达的度量值。
- 4-毒性逆转（poison reverse）：当拓扑变化时，路由器会将失效的路由标记为possibly down状态，并分配一个不可达的度量值。
- 5-抑制计时器（hold-down timer）：防止路由表频繁翻动，增加了网络的稳定性。
- 6-触发更新（trigger update）：一旦检测到路由崩溃，立即广播路由刷新报文，而不等等到下一刷新周期。





**THANK
YOU!**

