



河海大学 计算机与信息学院

计算机专业课程

计算机网络

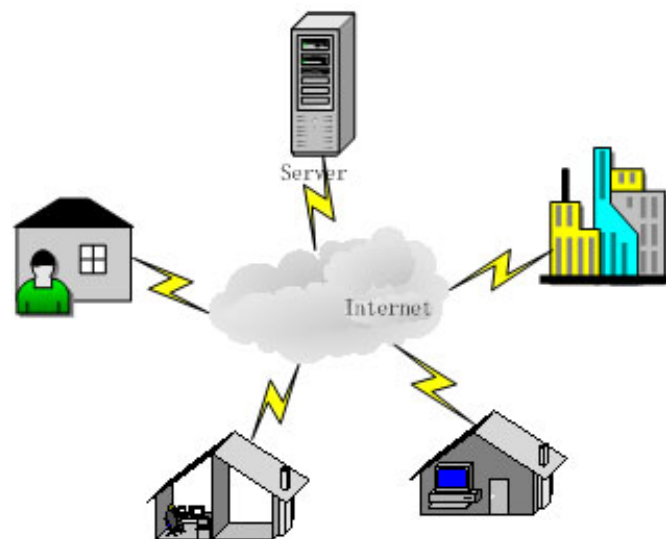
河海大学计算机与信息学院

2019年4月29日星期一



计算机专业课程

- 第1章 网络概述
- 第2章 网络体系结构
- 第3章 物理层
- 第4章 数据链路层
- 第5章 局域网
- 第6章 网络层**
- 第7章 传输层
- 第8章 应用层
- 第9章 网络管理和安全





因特网的路由选择协议



分层次的路由选择协议

- 因特网采用分层次的路由选择协议。
- 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。



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。



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。目前这类路由选择协议使用得最多，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。这样的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)



因特网的路由选择协议

- 内部网关协议 IGP: 具体的协议有多种, 如 RIP 和 OSPF 等。
- 外部网关协议 EGP: 目前使用的协议就是 BGP。



内部网关协议 RIP (Routing Information Protocol)

1. 工作原理

- ❑ 路由信息协议 **RIP** 是内部网关协议 **IGP**中最先得到广泛使用的协议。
- ❑ **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。



“距离”的定义

- 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 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 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”。



“距离”的定义

- RIP 允许一条路径最多只能包含 15 个路由器。
- “距离”的最大值为16 时即相当于不可达。可见 RIP 只适用于小型互联网。
- 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 协议的三个要点

- 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。



路由表的建立

- ❑ 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。

距离向量算法-如何根据RIP更新路由表

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不在路由表中，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，则把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，则进行更新，
否则，什么也不做。

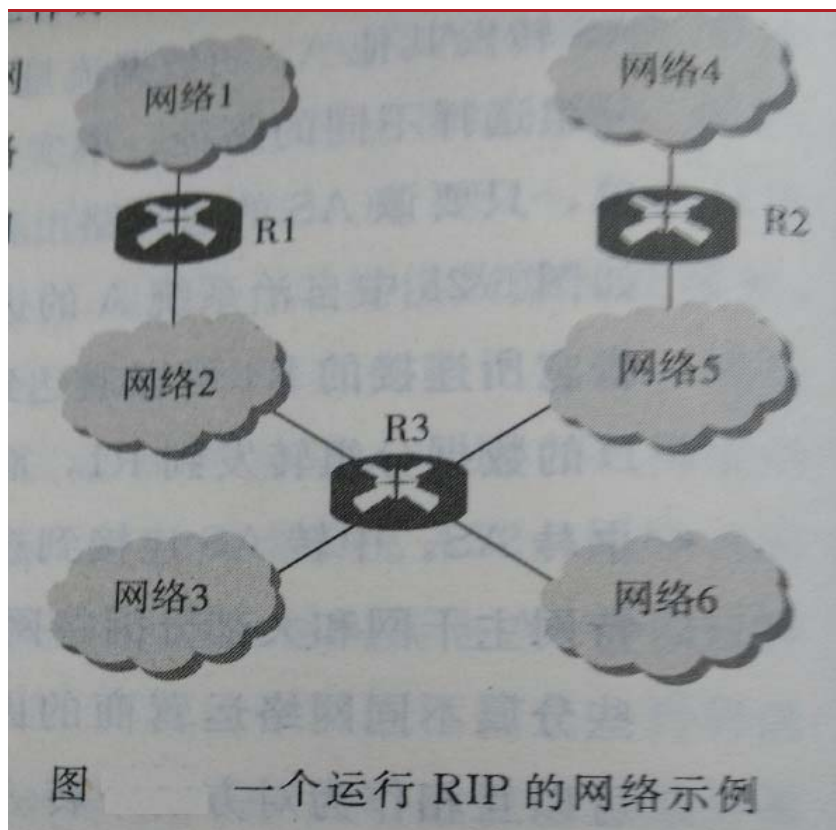
(3) 若 3 分钟还没有收到相邻路由器的更新路由表，则把此相邻路由器记为不可达路由器，即将距离置为16（距离为16表示不可达）。

(4) 返回。



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，它们的路由表当然也应当是不同的。

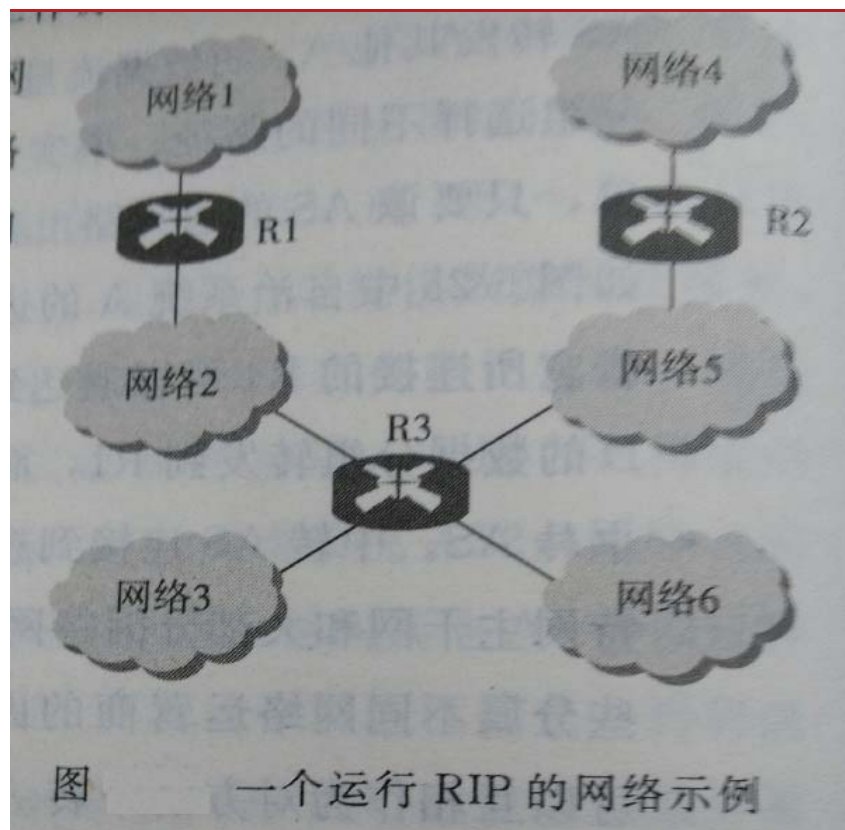


路由器R1转发表

目的子网	下一跳路由器	到目的子网跳数
子网1	R1	1
子网2	R1	1
子网3	—	∞
子网4	—	∞
子网5	—	∞
子网6	—	∞

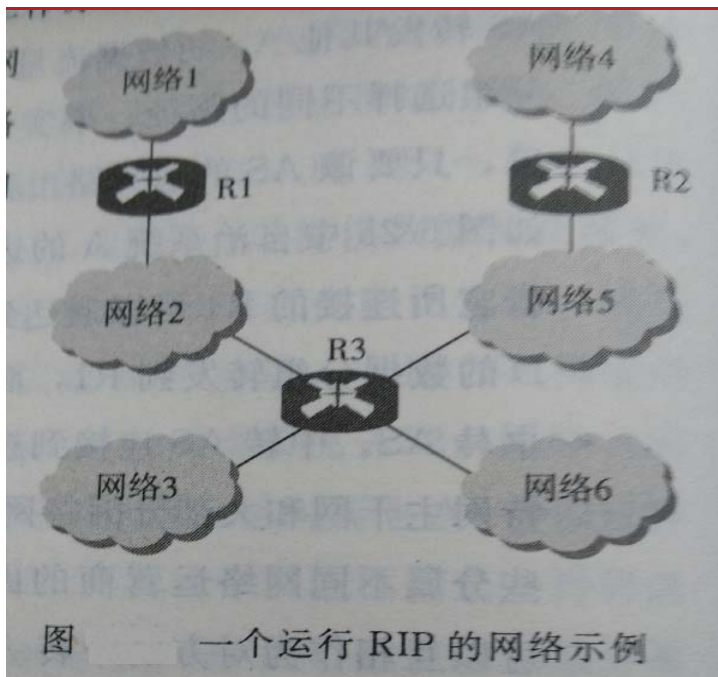
路由器R2转发表

目的子网	下一跳路由器	到目的子网跳数
子网1	—	∞
子网2	—	∞
子网3	—	∞
子网4	R2	1
子网5	R2	1
子网6	—	∞



路由器R3转发表

目的子网	下一跳路由器	到目的子网跳数
子网1	—	∞
子网2	R3	1
子网3	R3	1
子网4	—	∞
子网5	R3	1
子网6	R3	1



路由器交换一次

路由器R1转发表

目的子网	下一个路由器	到目的子网跳数
子网1	R1	1
子网2	R1	1
子网3	R3	2
子网4	—	∞
子网5	R3	2
子网6	R3	2

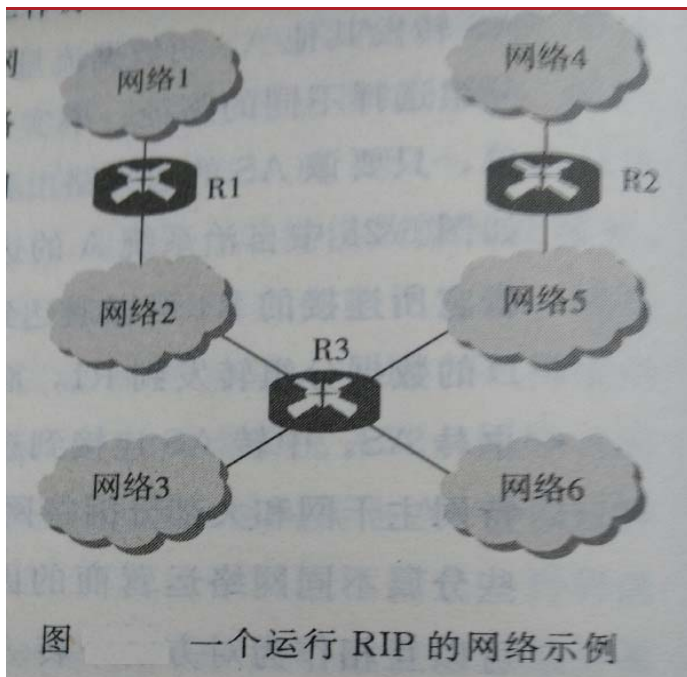
路由器R2转发表

目的子网	下一个路由器	到目的子网跳数
子网1	—	∞
子网2	R3	2
子网3	R3	2
子网4	R2	1
子网5	R2	1
子网6	R3	2

路由器R3转发表

目的子网	下一个路由器	到目的子网跳数
子网1	R1	2
子网2	R3	1
子网3	R3	1
子网4	R2	2
子网5	R3	1
子网6	R3	1

a) 一次迭代以后的路由表



路由器R1转发表

目的子网	下一个路由器	到目的子网跳数
子网1	R1	1
子网2	R1	1
子网3	R3	2
子网4	R3	3
子网5	R3	2
子网6	R3	2

路由器R2转发表

目的子网	下一个路由器	到目的子网跳数
子网1	R3	3
子网2	R3	2
子网3	R3	2
子网4	R2	1
子网5	R2	1
子网6	R3	2

路由器R3转发表

目的子网	下一个路由器	到目的子网跳数
子网1	R1	2
子网2	R3	1
子网3	R3	1
子网4	R2	2
子网5	R3	1
子网6	R3	1

b) 二次迭代以后的路由表

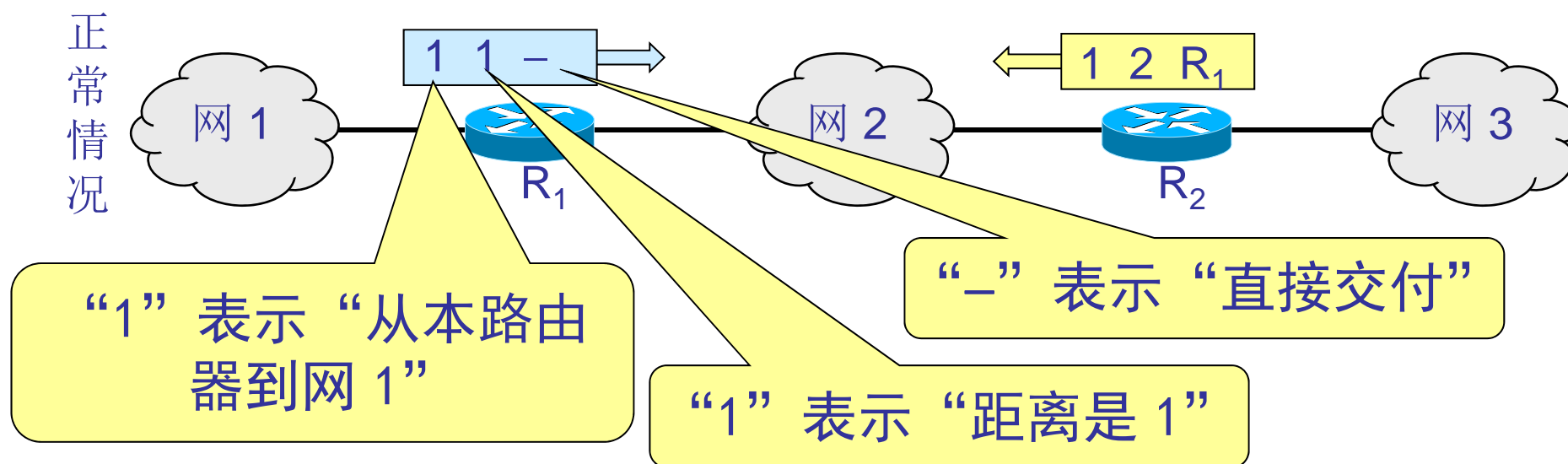


RIP 协议的优缺点

- ❑ RIP 存在的一个问题是当网络出现故障时，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 协议最大的优点就是实现简单，开销较小。
- ❑ 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。



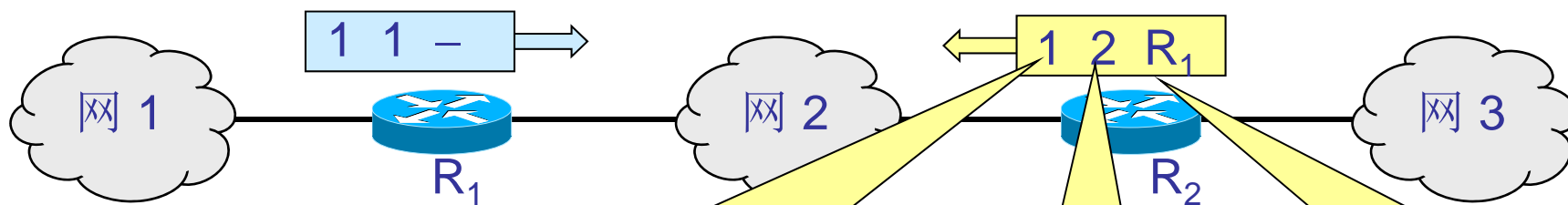
RIP 协议的缺点-无穷大问题



R_1 说：“我到网 1 的距离是 1，是直接交付。”



正常情况



“1” 表示 “从本路由器到网 1”

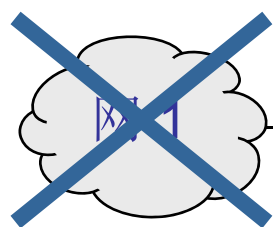
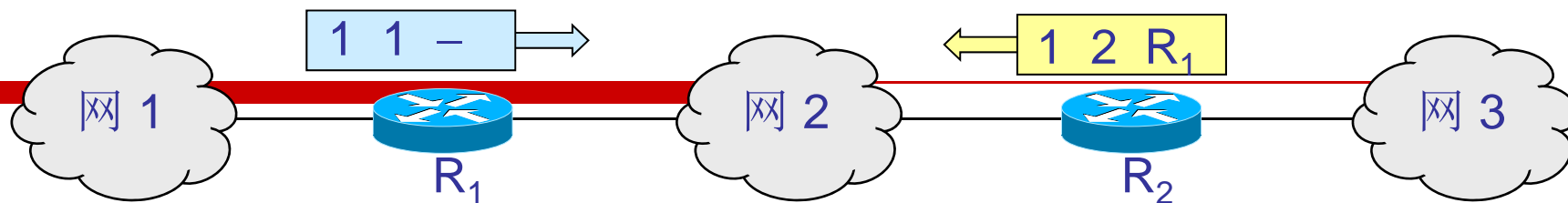
“ R_1 ” 表示
经过 R_1

“2” 表示 “距离是 2”

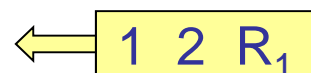
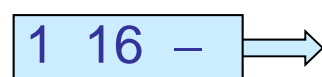
R_2 说: “我到网 1 的距离是 2, 是经过 R_1 。”



正常情况



网 1 出了故障

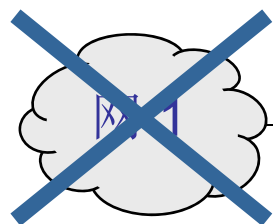
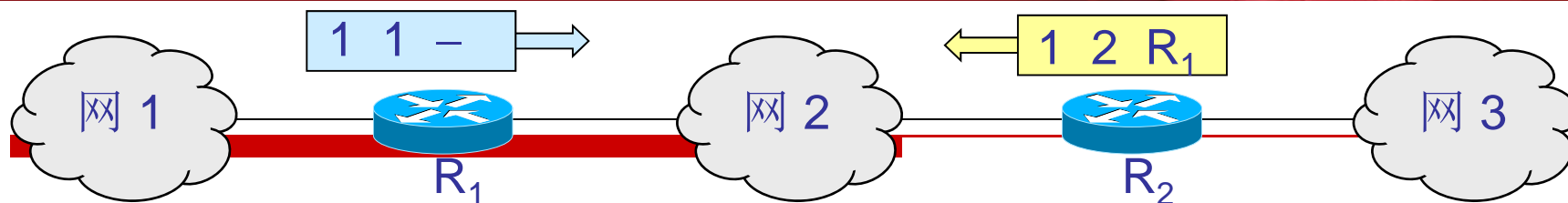


R_1 说：“我到网 1 的距离是 16（表示无法到达），是直接交付。”

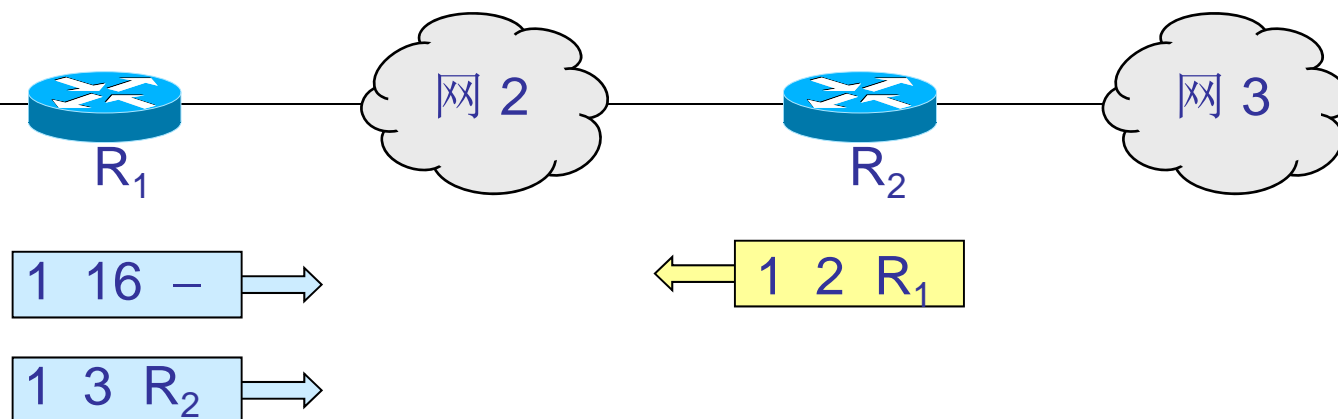
但 R_2 在收到 R_1 的更新报文之前，还发送原来的报文，因为这时 R_2 并不知道 R_1 出了故障。



正常情况



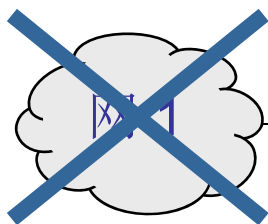
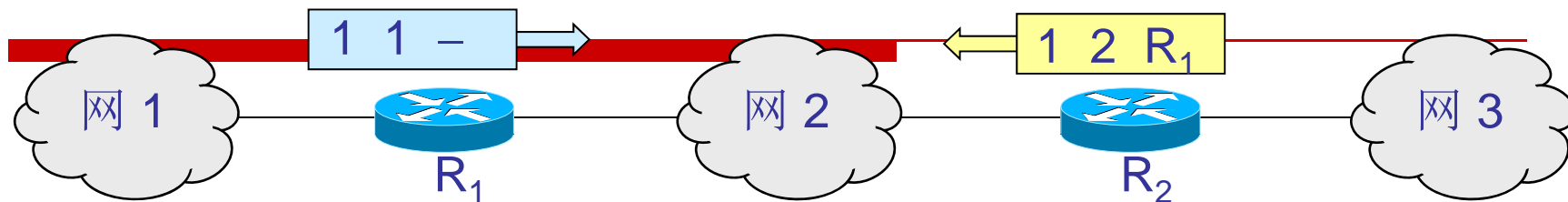
网 1 出了故障



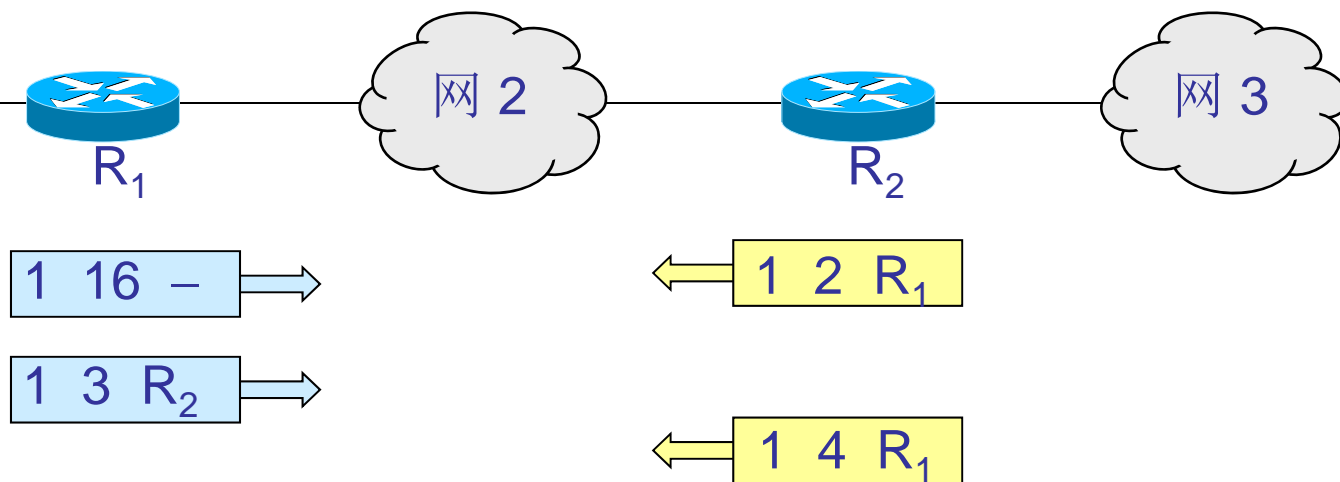
R_1 收到 R_2 的更新报文后，误认为可经过 R_2 到达网 1，于是更新自己的路由表，说：“我到网 1 的距离是 3，下一跳经过 R_2 ”。然后将此更新信息发送给 R_2 。



正常情况

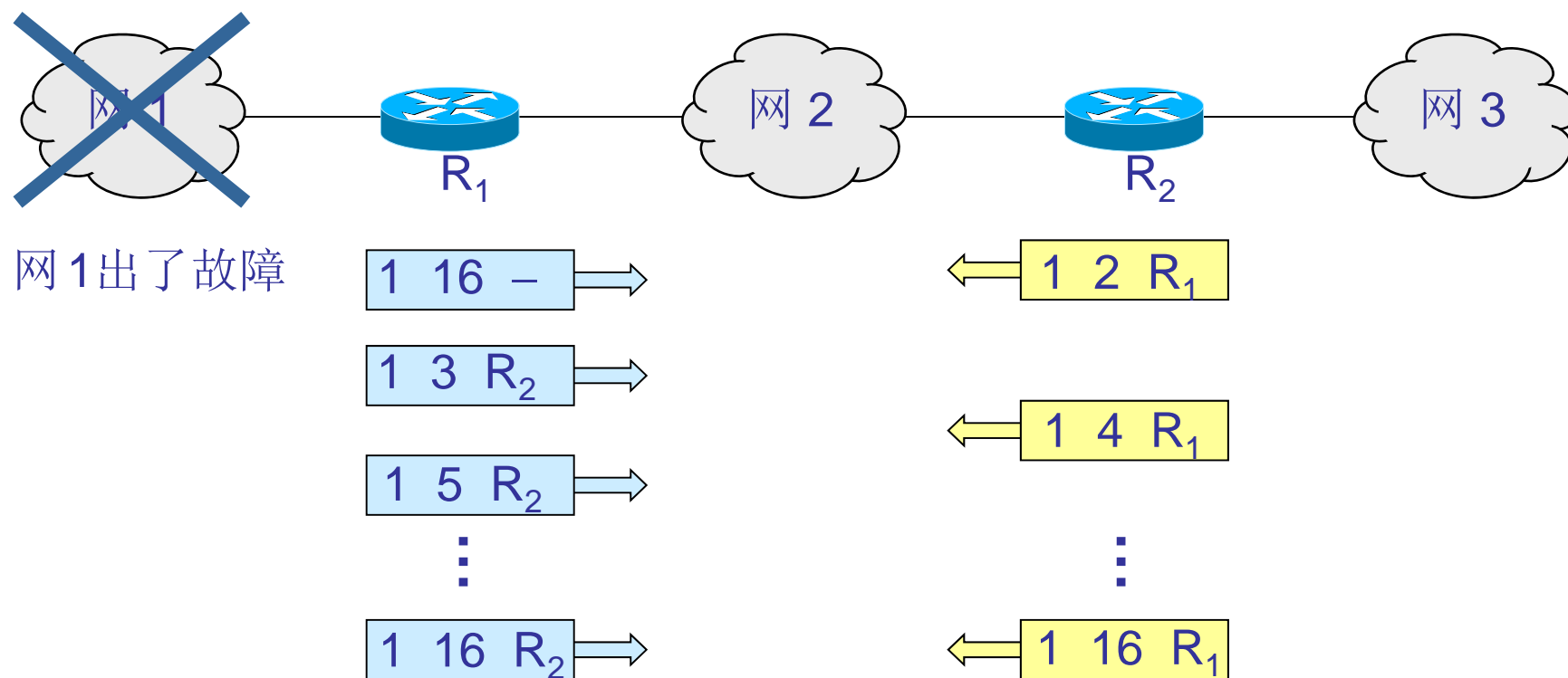


网 1 出了故障



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。



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。



内部网关协议 OSPF (Open Shortest Path First)

OSPF 协议的基本特点

- ❑ “开放”表明 OSPF 协议不是受某一家厂商控制，而是公开发表的。
- ❑ “最短路径优先”是因为使用了 Dijkstra 提出的最短路径算法 SPF
- ❑ OSPF 只是一个协议的名字，它并不表示其他的路由选择协议不是“最短路径优先”。
- ❑ 是分布式的链路状态协议。



三个要点

(1) 向本自治系统中所有路由器发送信息，这里使用的方法是洪泛法。

(2) 发送的信息就是与本路由器相邻的所有路由器的链路状态。

■ “链路状态”就是说明本路由器都和哪些路由器相邻，以及该链路的“度量”(metric)。

(3) 只有当链路状态发生变化时，路由器才用洪泛法向所有路由器发送此信息。



链路状态数据库(link-state database)

- 由于各路由器之间频繁地交换链路状态信息，因此所有的路由器最终都能建立一个链路状态数据库。
- 这个数据库实际上就是全网的拓扑结构图，它在全网范围内是一致的（这称为链路状态数据库的同步）。
- OSPF 的链路状态数据库能较快地进行更新，使各个路由器能及时更新其路由表。OSPF 的更新过程收敛得快是其重要优点。



外部网关协议 BGP

- ❑ BGP 是不同自治系统的路由器之间交换路由信息的协议。
- ❑ BGP 较新版本是 2006 年 1 月发表的 BGP-4（BGP 第 4 个版本），即 RFC 4271 ~ 4278。
- ❑ 可以将 BGP-4 简写为 BGP。



BGP 使用的环境

- 因特网的规模太大，使得自治系统之间路由选择非常困难。对于自治系统之间的路由选择，要寻找最佳路由是很不现实的。
 - 当一条路径通过几个不同 AS 时，要想对这样的路径计算出有意义的代价是不太可能的。
 - 比较合理的做法是在 AS 之间交换“可达性”信息。
- 自治系统之间的路由选择必须考虑有关策略。
- 因此，边界网关协议 BGP 只能是力求寻找一条能够到达目的网络且**比较好的路由**（不能兜圈子），而**并非要寻找一条最佳路由**。



BGP 发言人(BGP speaker)

- 每一个自治系统的管理员要选择至少一个路由器作为该自治系统的“**BGP 发言人**”。
- 一般说来，两个 BGP 发言人都是通过一个共享网络连接在一起的，而 BGP 发言人往往就是 BGP 边界路由器。

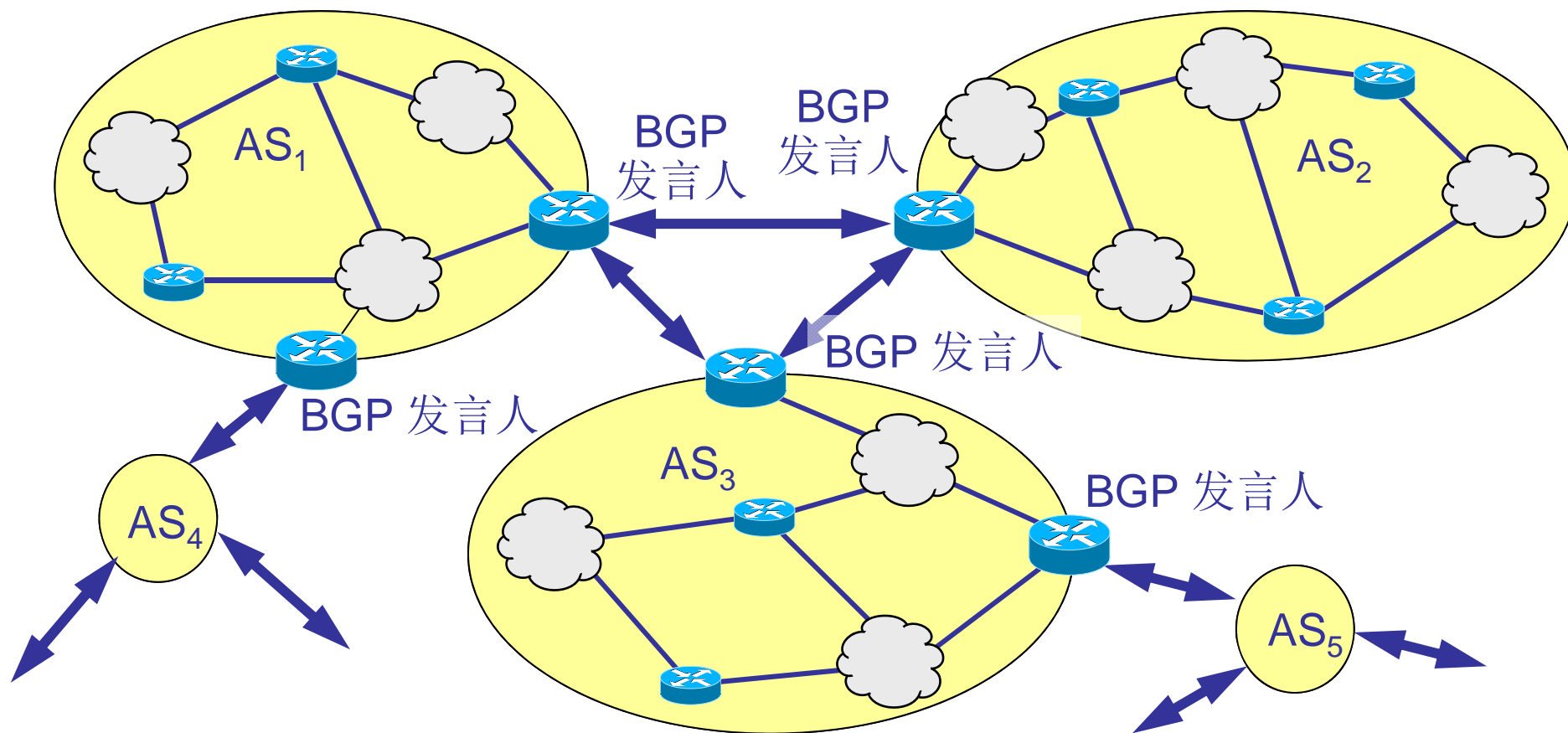


BGP 交换路由信息

- 一个 BGP 发言人与其他自治系统中的 BGP 发言人要交换路由信息，就要先建立 TCP 连接，然后在此连接上交换 BGP 报文以建立 BGP 会话(session)，利用 BGP 会话交换路由信息。
- 使用 TCP 连接能提供可靠的服务，也简化了路由选择协议。
- 使用 TCP 连接交换路由信息的两个 BGP 发言人，彼此成为对方的邻站或对等站。



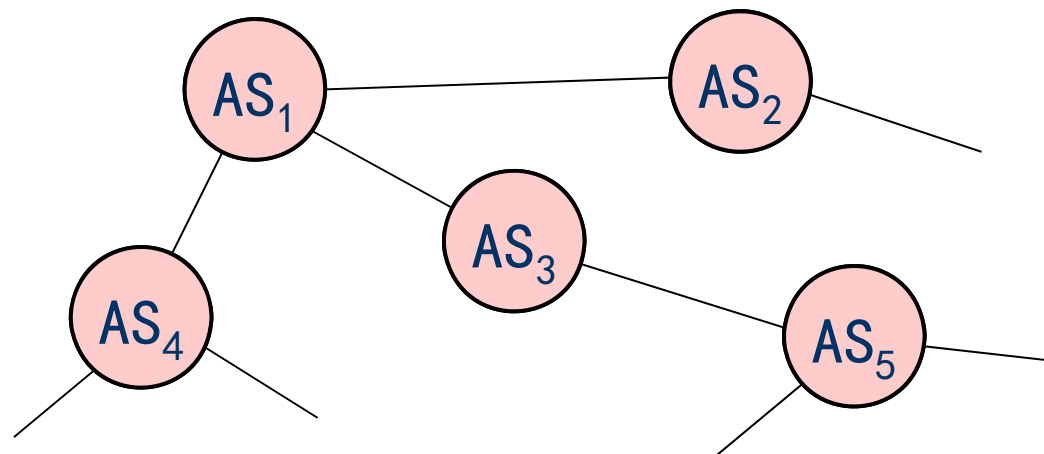
BGP 发言人和自治系统 AS 的关系





AS 的连通图举例

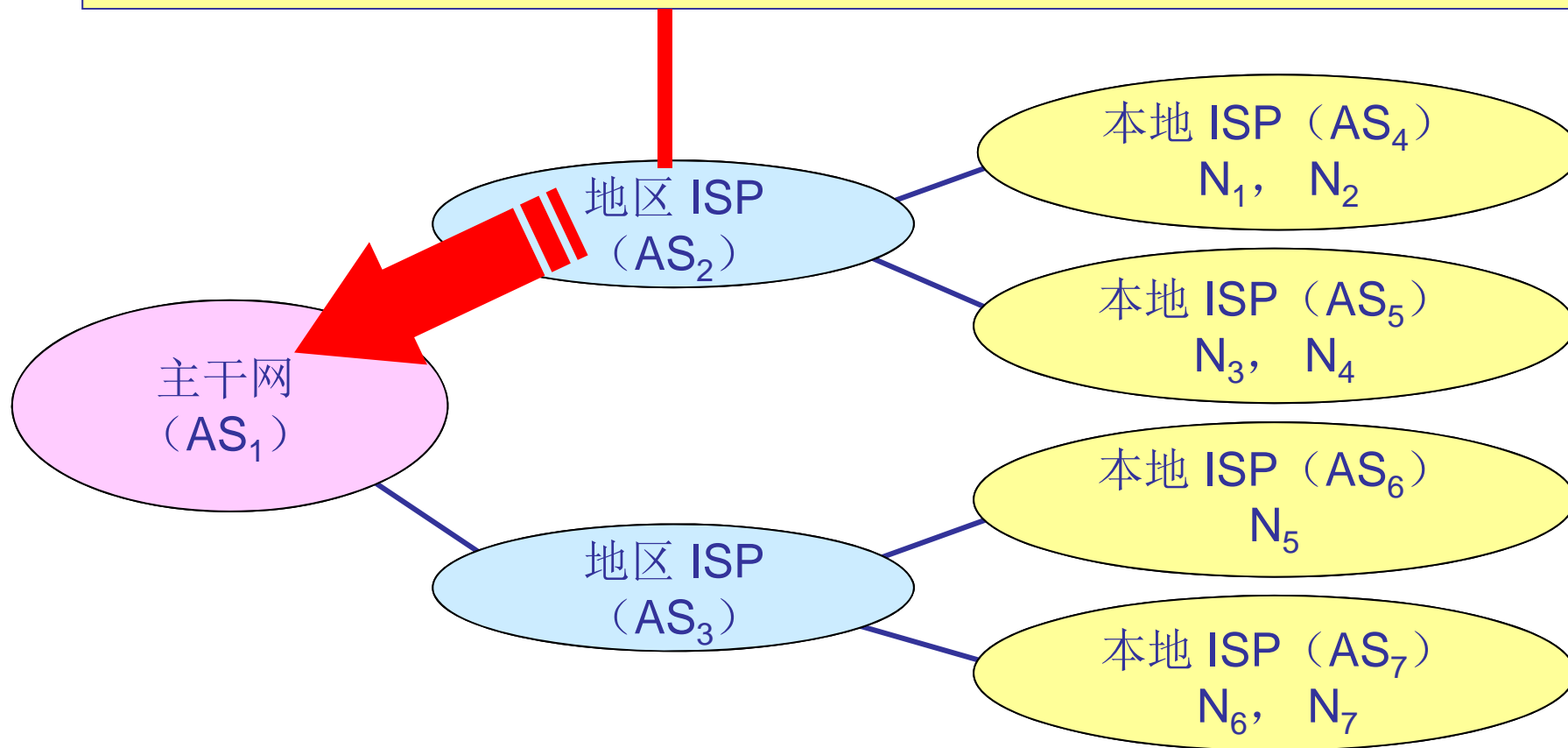
- BGP 所交换的网络可达性的信息就是要到达某个网络所要经过的一系列 AS。
- 当 BGP 发言人互相交换了网络可达性的信息后，各 BGP 发言人就根据所采用的策略从收到的路由信息中找出到达各 AS 的较好路由。





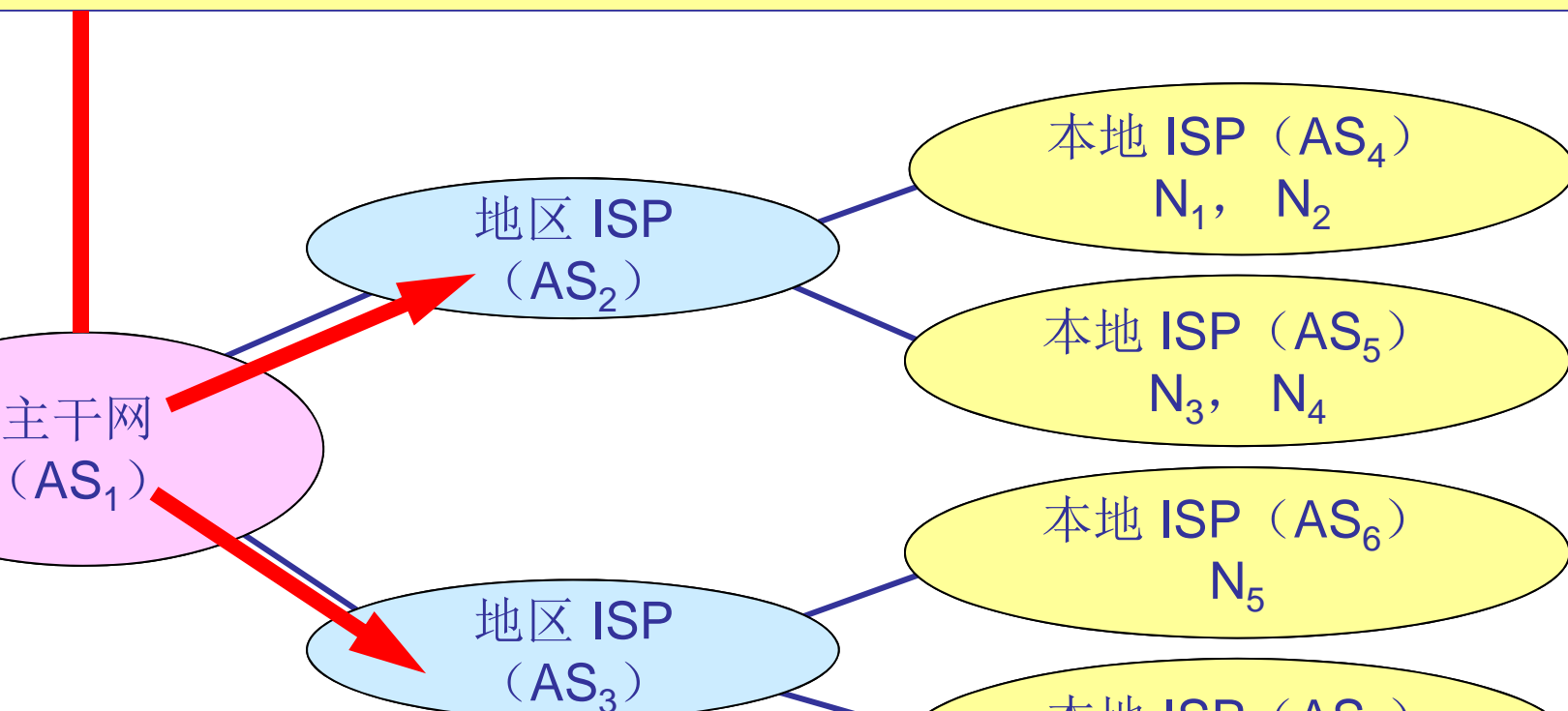
BGP 发言人交换路径向量

自治系统 AS_2 的 BGP 发言人通知主干网的 BGP 发言人：
“要到达网络 N_1, N_2, N_3 和 N_4 可经过 AS_2 。”



GP 发言人交换路径向量

主干网还可发出通知：“要到达网络 N_5 , N_6 和 N_7 可沿路径 (AS_1, AS_3) 。”

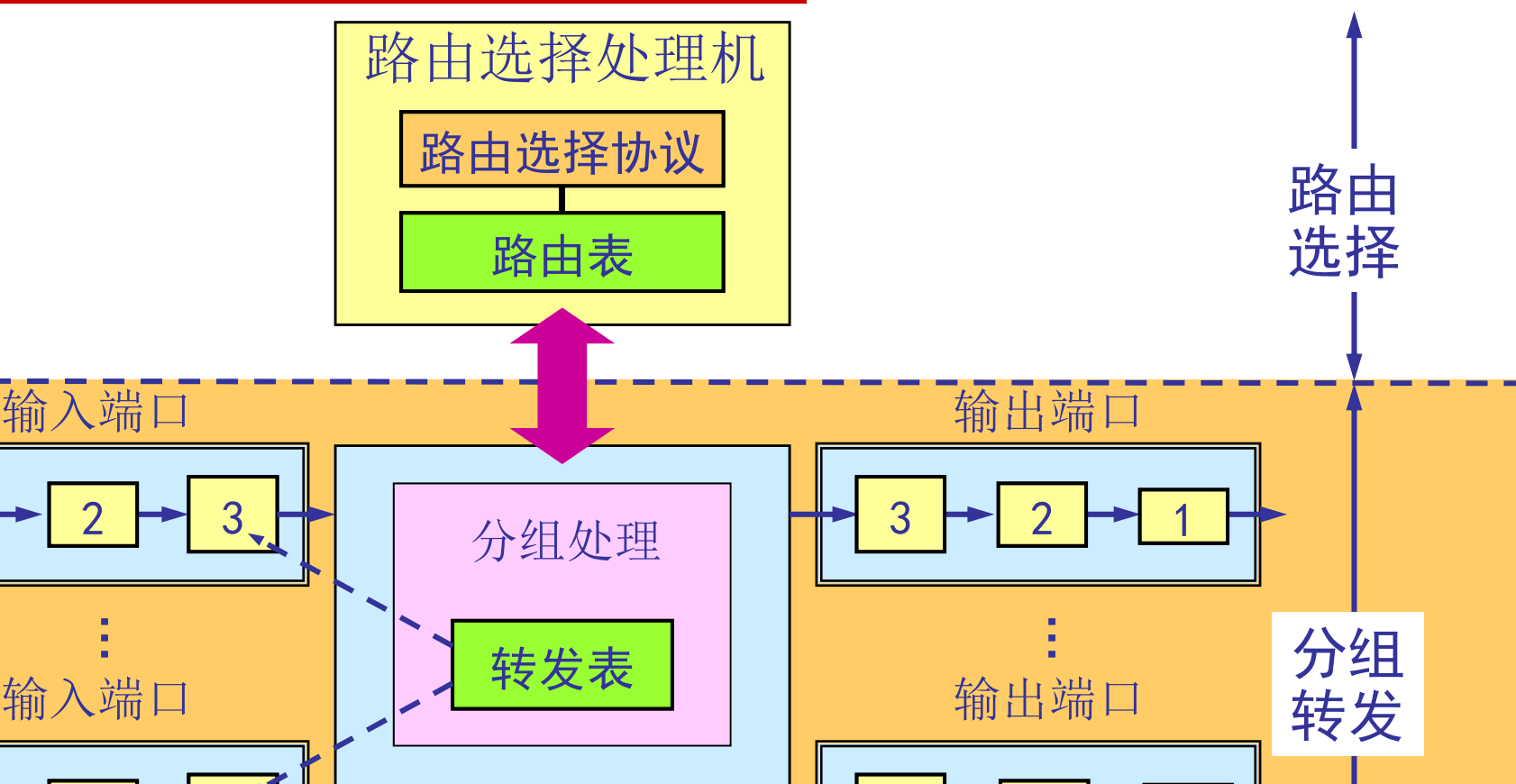


由器

路由器是一种具有多个输入端口和多个输出端口的专用计算机，其任务是转发分组。也就是说，将路由器某个输入端口收到的分组，按照分组要去的目的地（即目的网络），把该分组从路由器的某个合适的输出端口转发给下一跳路由器。下一跳路由器也按照这种方法处理分组，直到该分组到达终点为止。

路由器的两个主要功能：路由选择和分组转发。

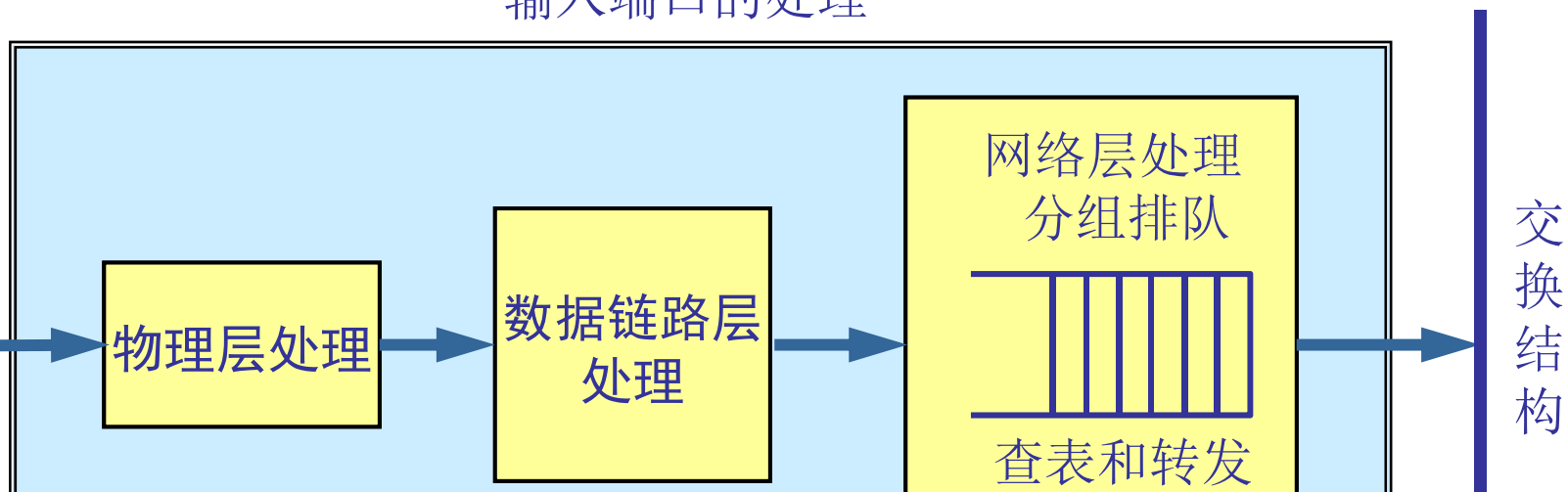
型的路由器的结构



入端口分组的处理

数据链路层剥去帧首部和尾部后，将分组送到网络层的队列中排队等待处理。这会产生一定的时延。

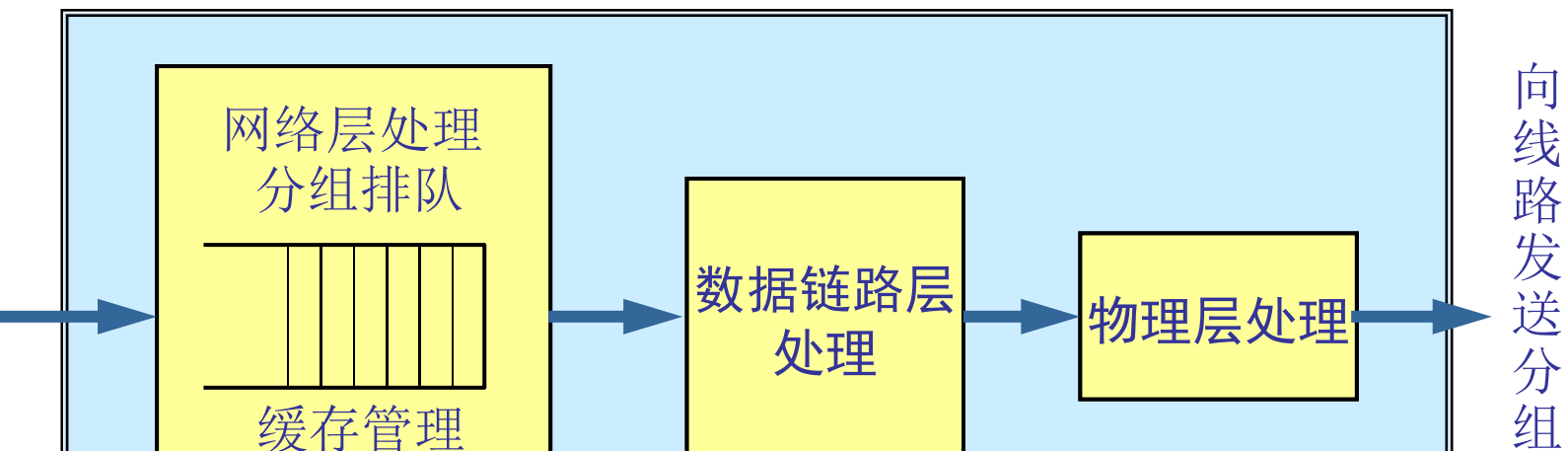
输入端口的处理



出端口对分组的处理

当交换结构传送过来的分组先进行缓存。数据链路层处理模块将分组加上链路层的首部和尾部，交给物理层后发送到外部线路。

输出端口的处理



组丢弃

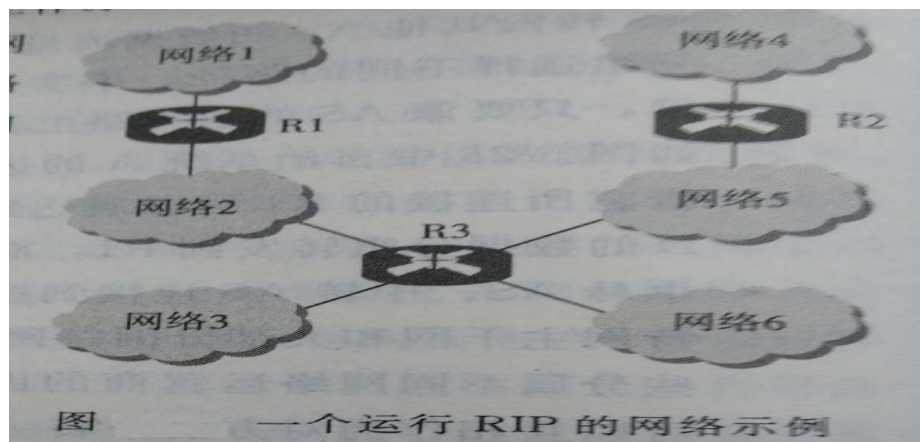
若路由器处理分组的速率赶不上分组进入队列的速率，则队列的存储空间最终必定减少到零，这就使后面再进入队列的分组由于没有存储空间而只能被丢弃。

路由器中的输入或输出队列产生溢出是造成分组丢失的重要原因。

作业

简述什么是自治系统？什么是IGP,EGP,BGP.

简述RIP协议的工作过程。有如下网络拓扑结构图，



写出初始状态各路由器的路由表，以及经过一次转发后路由表。