

本节内容

路由算法及路由协议

王道考研/CSKAOYAN.COM

1

路由算法

R1

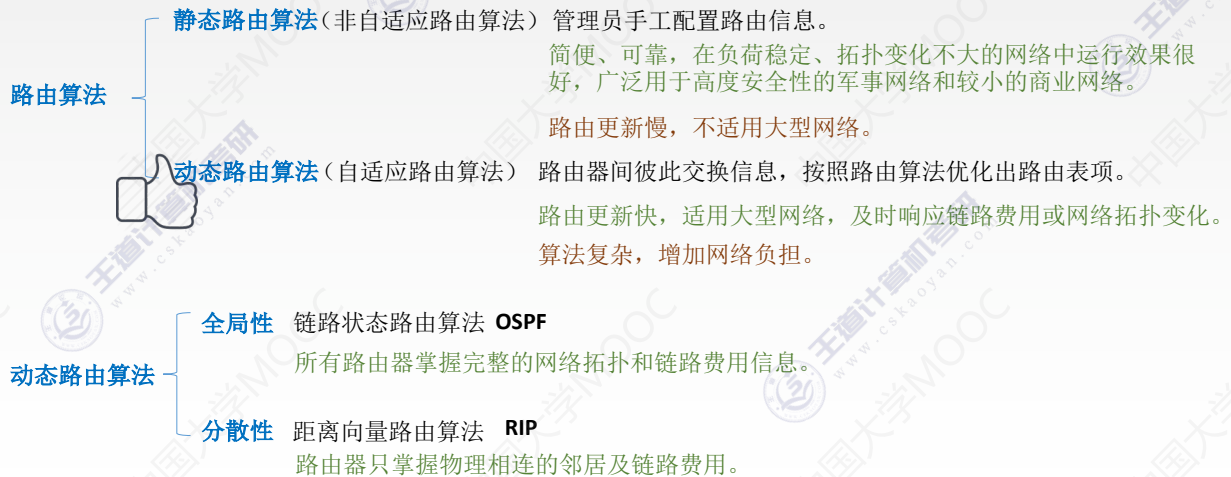
目的网络IP地址	子网掩码	下一跳IP地址	接口

最佳路由：“最佳”只能是相对于某一种特定要求下得出的较为合理的选择而已。

王道考研/CSKAOYAN.COM

2

路由算法的分类



王道考研/CSKAOYAN.COM

3

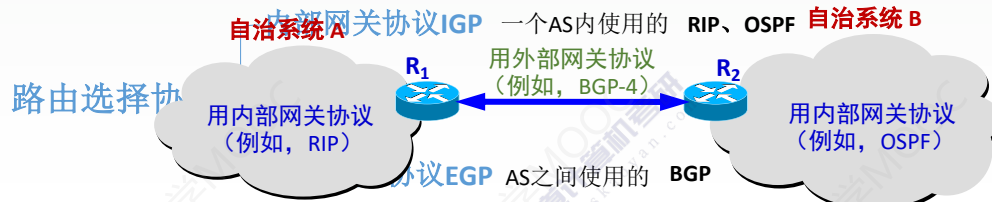
分层次的路由选择协议

- (1) 因特网规模很大
- (2) 许多单位不想让外界知道自己的路由选择协议，但还想连入因特网



自治系统AS：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之间的路由协议以确定在AS之间的路由。

一个AS内的所有网络都属于一个行政单位来管辖，一个自治系统的所有路由器在本自治系统内都必须连通。



王道考研/CSKAOYAN.COM

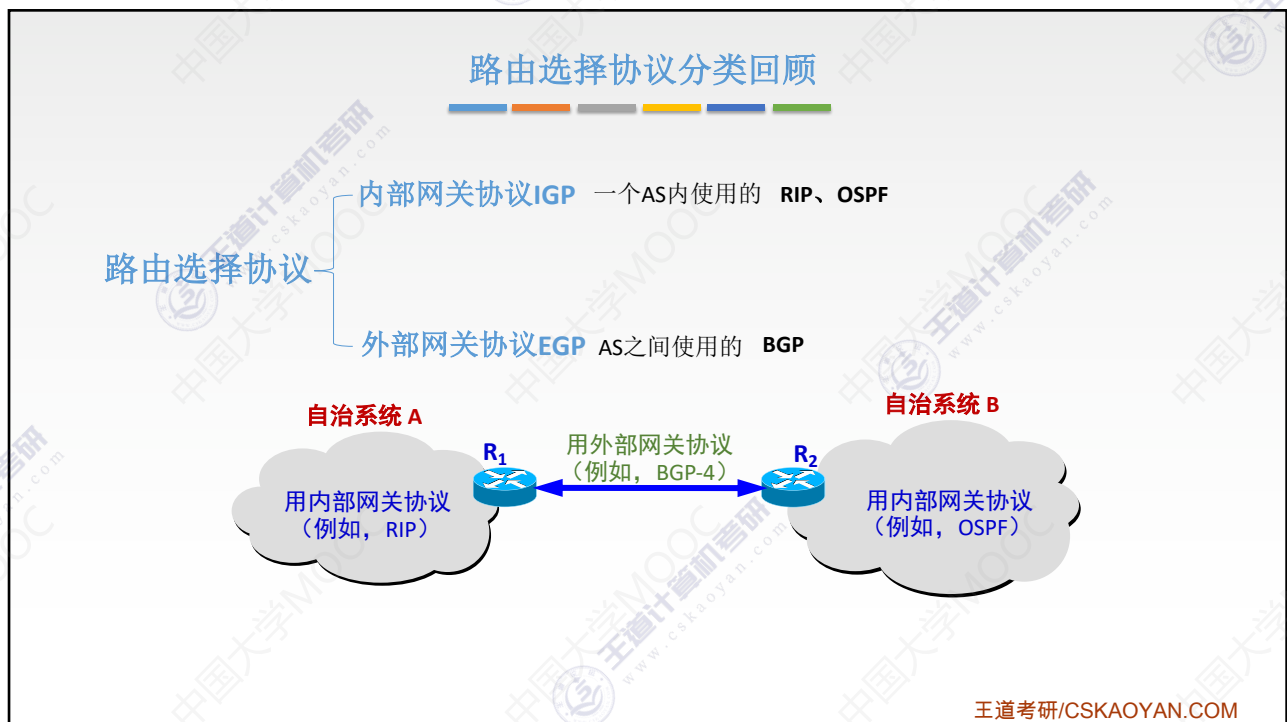
4

本节内容

RIP协议及距离向量算法

王道考研/CSKAOYAN.COM

5



6

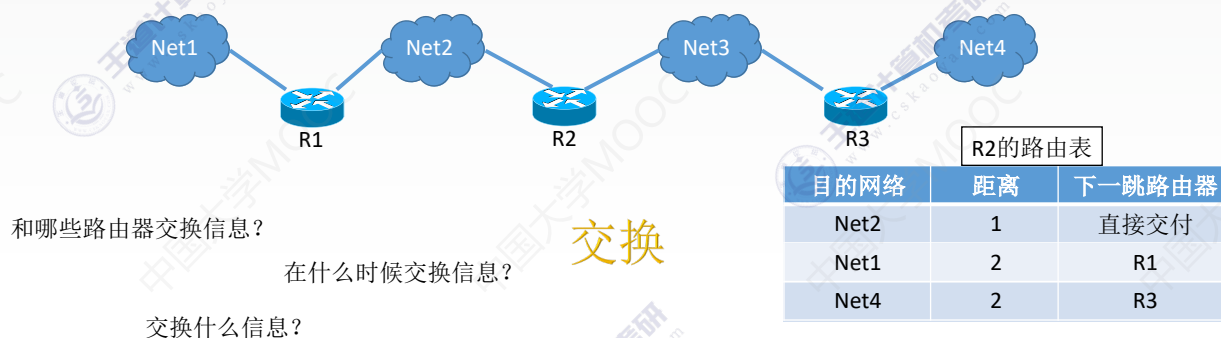
RIP协议

RIP是一种分布式的基于**距离向量**的路由选择协议，是因特网的协议标准，最大优点是**简单**。

RIP协议要求网络中每一个路由器都维护**从它自己到其他每一个目的网络的唯一最佳距离记录**（即一组距离）。

距离：通常为“跳数”，即从源端口到目的端口所经过的路由器个数，经过一个路由器跳数+1。特别的，从一路由器到直接连接的网络距离为1。RIP允许一条路由最多只能包含15个路由器，因此距离为**16表示网络不可达**。

➡ RIP协议只适用于**小互联网**。



王道考研/CSKAOYAN.COM

7

RIP协议和谁交换? 多久交换一次? 交换什么?

1. 仅和**相邻路由器**交换信息。

2. 路由器交换的信息是**自己的路由表**。



“我到Net1网络的（最短）距离是5跳，下一跳应该走R1路由器……”

3. **每30秒**交换一次路由信息，然后路由器根据新信息更新路由表。若超过**180s**没收到邻居路由器的通告，则判定邻居没了，并更新自己路由表。

路由器刚开始工作时，只知道直接连接的网络的距离（距离为1），接着每一个路由器也只会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，所有路由器最终都会知道到达本自治系统任何一个网络的最短距离和下一跳路由器的地址，即“**收敛**”。



路由表怎么更新的?

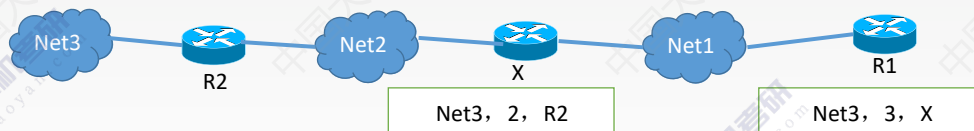
王道考研/CSKAOYAN.COM

8

距离向量算法

1. 修改相邻路由器发来的RIP报文中所有表项

对地址为X的相邻路由器发来的RIP报文，修改此报文中的所有项目：把“下一跳”字段中的地址改为X，并把所有的“距离”字段+1。



2. 对修改后的RIP报文中的每一个项目，进行以下步骤：

(1) R1路由表中若没有Net3，则把该项目填入R1路由表

(2) R1路由表中若有Net3，则查看下一跳路由器地址：

若下一跳是X，则用收到的项目替换源路由表中的项目；

若下一跳不是X，原来距离比从X走的距离远则更新，否则不作处理。

3. 若180s还没收到相邻路由器X的更新路由表，则把X记为不可达的路由器，即把距离设置为16。

4. 返回

王道考研/CSKAOYAN.COM

9

距离向量算法练习1

已知路由器R6的路由表，现收到相邻路由器R4发来的路由更新信息，试更新路由器R1的路由表：

R6的路由表

目的网络	距离	下一跳路由器
Net2	3	R4
Net3	4	R5
...

R4发来的路由更新信息

目的网络	距离	下一跳路由器
Net1	3	R1
Net2	4	R2
Net3	1	直接交付

解：

R4发来的路由更新信息-修改版

目的网络	距离	下一跳路由器
Net1	4	R4
Net2	5	R4
Net3	2	R4

R6更新后的路由表

目的网络	距离	下一跳路由器
Net1	4	R4
Net2	5	R4
Net3	2	R4
...

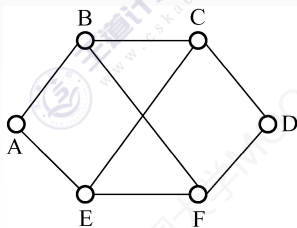
王道考研/CSKAOYAN.COM

10

距离向量算法练习2

考虑如图所示的子网，该子网使用了距离-向量算法，下面的向量刚刚到达路由器C：来自B的向量为（5，0，8，12，6，2）；来自D的向量为（16，12，6，0，9，10）；来自E的向量为（7，6，3，9，0，4）。经过测量，C到B、D和E的延迟分别为6，3和5，那么C到达所有结点的最短路径是（ ）。

- A. (5, 6, 0, 9, 6, 2) B. (11, 6, 0, 3, 5, 8)
C. (5, 11, 0, 12, 8, 9) D. (11, 8, 0, 7, 4, 9)



C到B: (11, 6, 14, 18, 12, 8)

C到D: (19, 15, 9, 3, 12, 13)

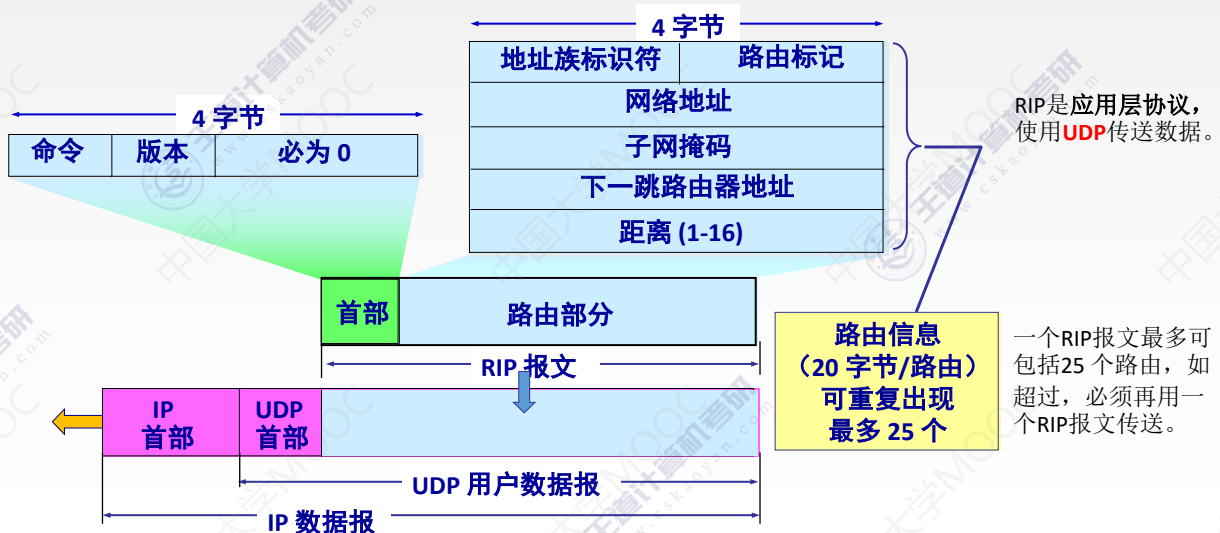
C到E: (12, 11, 8, 14, 5, 9)

(11, 6, 0, 3, 5, 8)

王道考研/CSKAOYAN.COM

11

RIP协议的报文格式

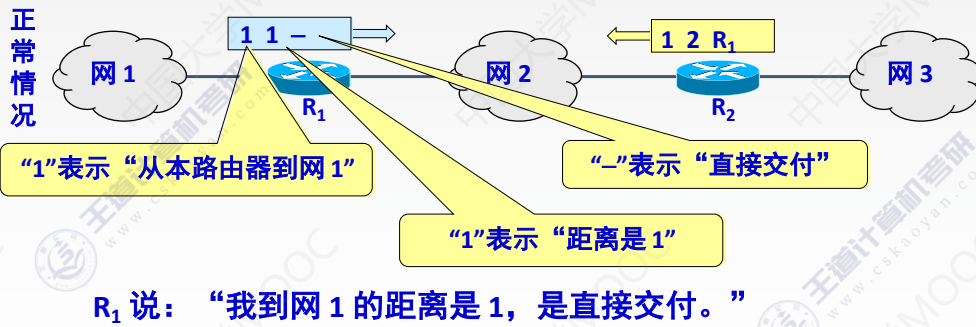


王道考研/CSKAOYAN.COM

12

RIP协议好消息传得快，坏消息传得慢

RIP的特点：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，“慢收敛”。

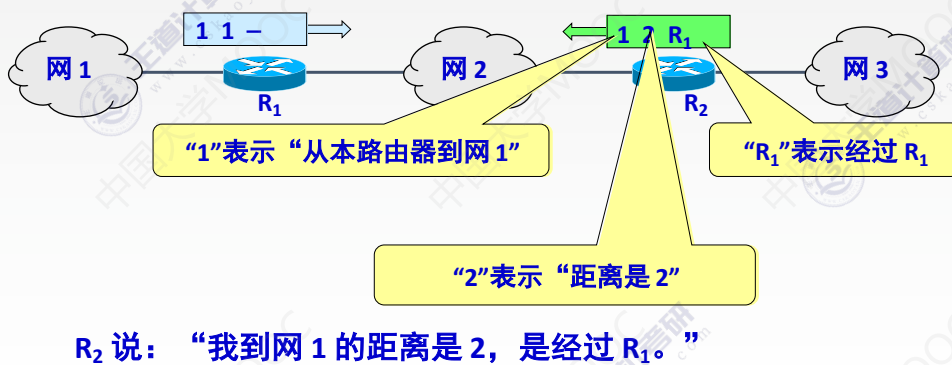


王道考研/CSKAOYAN.COM

13

RIP协议好消息传得快，坏消息传得慢

RIP的特点：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，“慢收敛”。



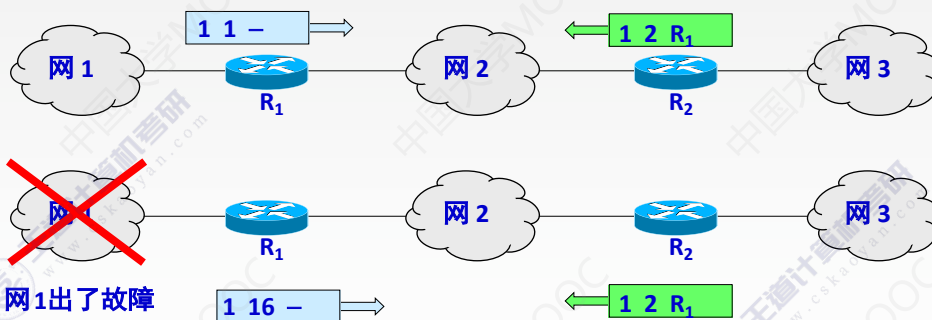
王道考研/CSKAOYAN.COM

14

RIP协议好消息传得快，坏消息传得慢

RIP的特点：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，“慢收敛”。

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。

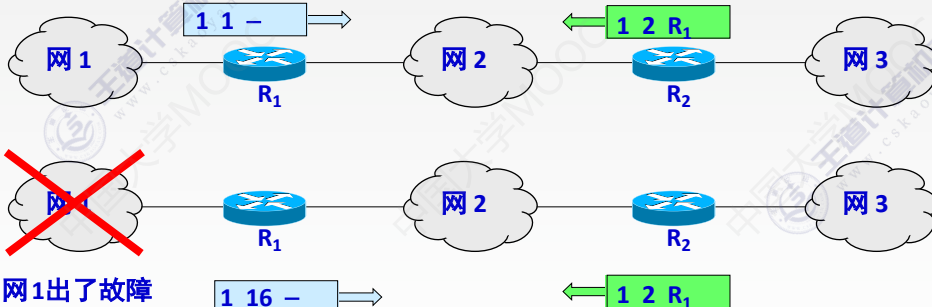
王道考研/CSKAOYAN.COM

15

RIP协议好消息传得快，坏消息传得慢

RIP的特点：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，“慢收敛”。

正常情况



网1出了故障

R₁ 收到 R₂ 的更新报文后，误认为可经过 R₂ 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 R₂”。然后将此更新信息发送给 R₂。

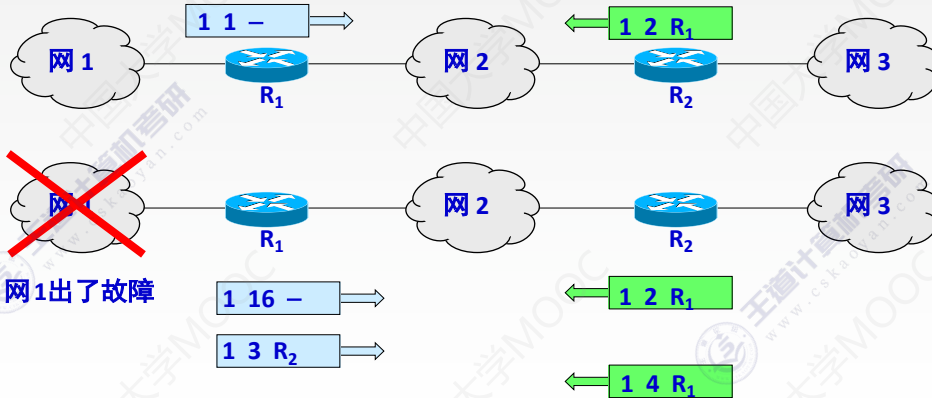
王道考研/CSKAOYAN.COM

16

RIP协议好消息传得快，坏消息传得慢

RIP的特点：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，“慢收敛”。

正常情况



网1出了故障

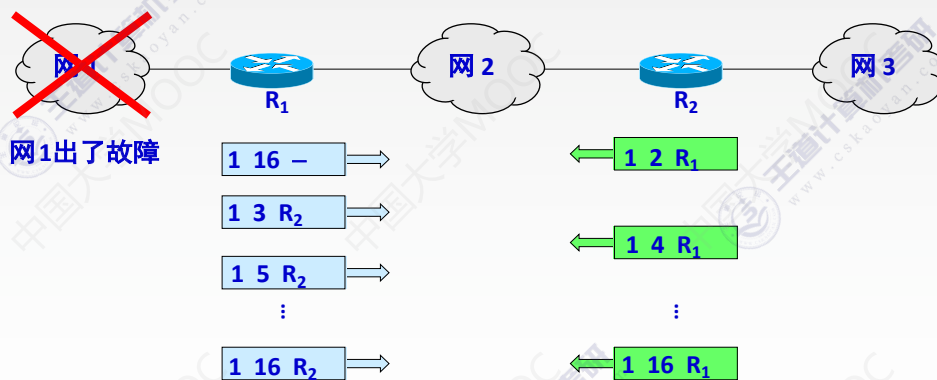


R₂ 以后又更新自己的路由表为“1, 4, R₁”，表明“我到网1距离是4，下一跳经过R₁”。

王道考研/CSKAOYAN.COM

17

RIP协议好消息传得快，坏消息传得慢



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。

王道考研/CSKAOYAN.COM

18

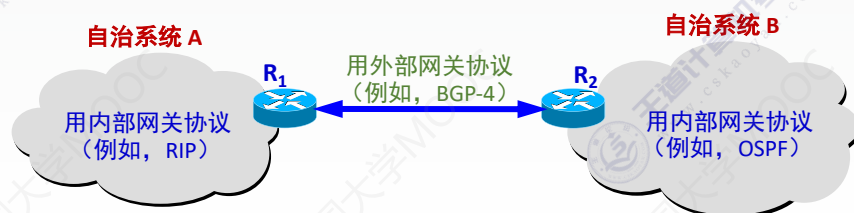


19



20

路由选择协议分类回顾



王道考研/CSKAOYAN.COM

21

OSPF协议

开放最短路径优先OSPF协议：“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最主要的特征就是使用分布式的链路状态协议。

OSPF的特点：

- | | |
|-------|--|
| 和谁交换？ | 1.使用洪泛法向自治系统内 所有路由器 发送信息，即路由器通过输出端口向所有相邻的路由器发送信息，而每一个相邻路由器又再次将此信息发往其所有的相邻路由器。 广播 |
| 交换什么？ | 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 （本路由器和哪些路由器相邻，以及该链路的度量/代价——费用、距离、时延、带宽等）。 |
| 多久交换？ | 3.只有当 链路状态发生变化 时，路由器才向所有路由器洪泛发送此信息。 |

最后，所有路由器都能建立一个**链路状态数据库**，即**全网拓扑图**。

王道考研/CSKAOYAN.COM

22

链路状态路由算法

1. 每个路由器发现它的邻居结点【HELLO问候分组】，并了解邻居节点的网络地址。
 2. 设置到它的每个邻居的成本度量metric。
 3. 构造【DD数据库描述分组】，向邻站给出自己的链路状态数据库中的所有链路状态项目的摘要信息。
 4. 如果DD分组中的摘要自己都有，则邻站不做处理；如果有没有的或者是更新的，则发送【LSR链路状态请求分组】请求自己没有的和比自己更新的信息。
 5. 收到邻站的LSR分组后，发送【LSU链路状态更新分组】进行更新。
 6. 更新完毕后，邻站返回一个【LSAck链路状态确认分组】进行确认。
- 只要一个路由器的链路状态发生变化：
5. 泛洪发送【LSU链路状态更新分组】进行更新。
 6. 更新完毕后，其他站返回一个【LSAck链路状态确认分组】进行确认。
 7. 使用Dijkstra根据自己的链路状态数据库构造到其他节点间的最短路径。

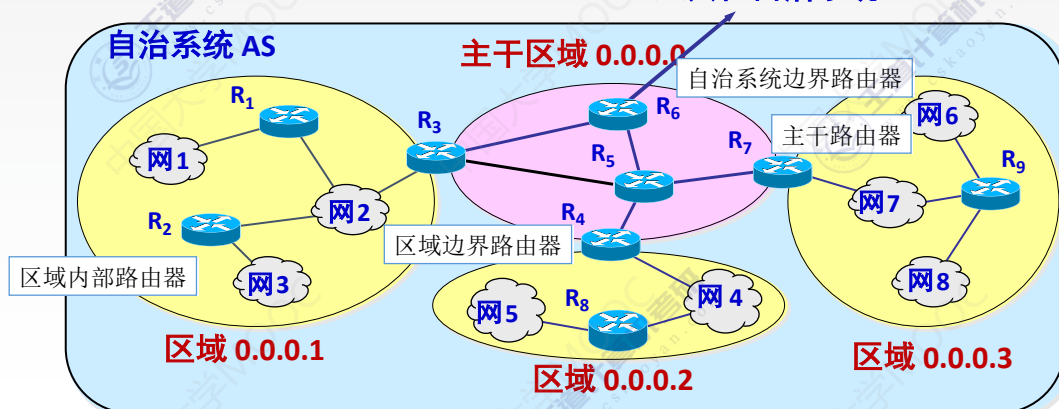
王道考研/CSKAOYAN.COM

23

OSPF的区域

为了使 OSPF 能够用于规模很大的网络，OSPF 将一个自治系统再划分为若干个更小的范围，叫做区域。每一个区域都有一个 32 位的区域标识符（用点分十进制表示）。区域也不能太大，在一个区域内的路由器最好不要超过 200 个。

至其他自治系统



王道考研/CSKAOYAN.COM

24



25

OSPF其他特点

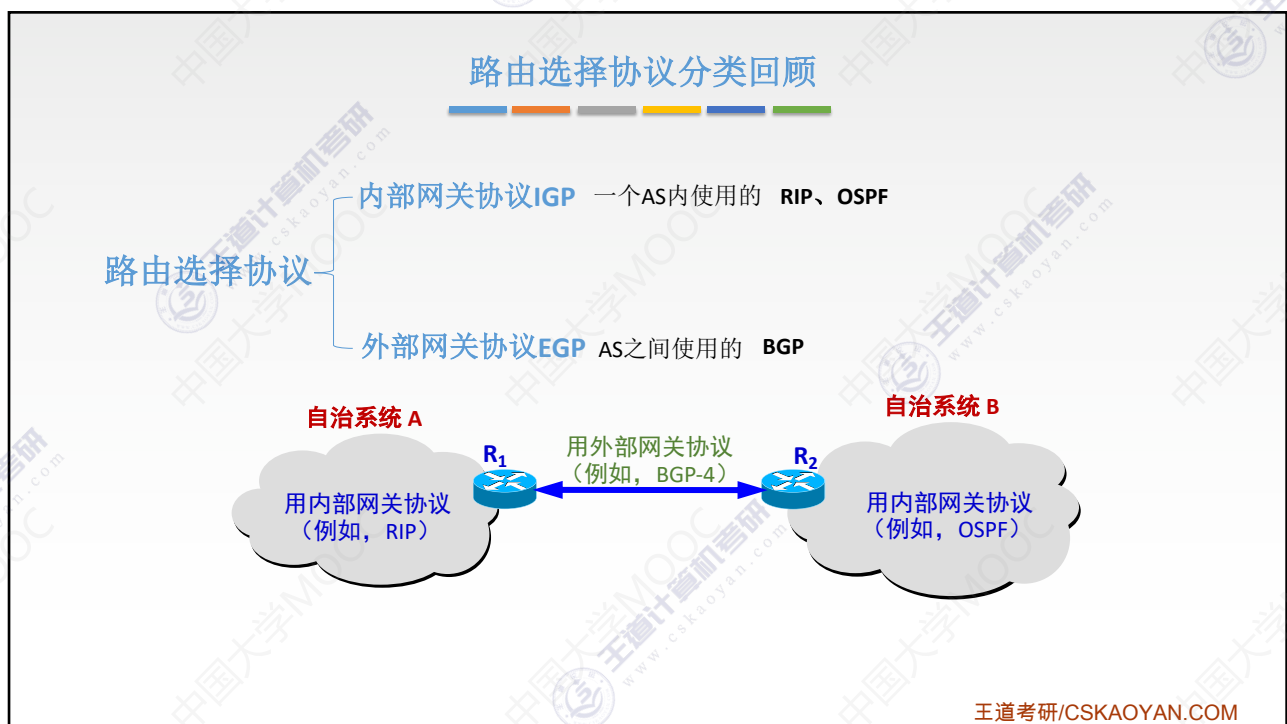
- 1.每隔**30min**，要刷新一次数据库中的链路状态。
- 2.由于一个路由器的链路状态只涉及到与相邻路由器的连通状态，因而与整个互联网的规模并无直接关系。因此当**互联网规模很大**时，OSPF 协议要比距离向量协议 RIP 好得多。
- 3.OSPF不存在坏消息传的慢的问题，它的**收敛速度很快**。

王道考研/CSKAOYAN.COM

26



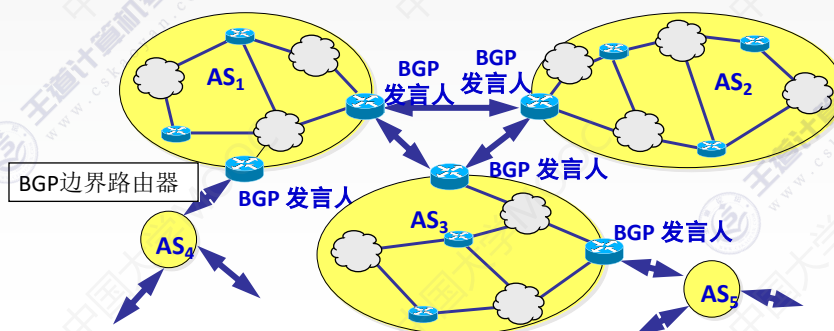
27



28

BGP协议

- 和谁交换？ 与其他AS的邻站BGP发言人交换信息。
- 交换什么？ 交换的网络可达性的信息，即要到达某个网络所要经过的一系列AS。
- 多久交换？ 发生变化时更新有变化的部分。

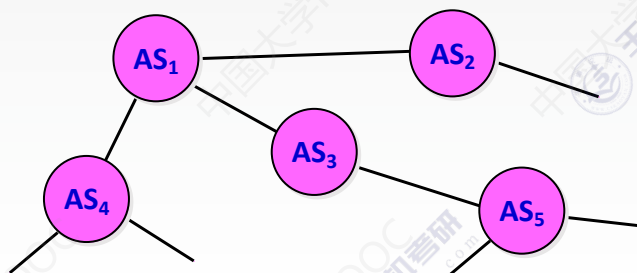


王道考研/CSKAOYAN.COM

29

BGP协议交换信息的过程

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。



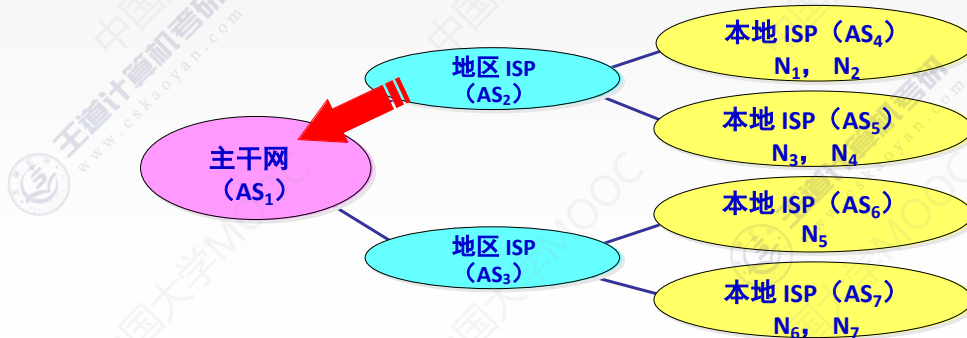
王道考研/CSKAOYAN.COM

30

BGP协议交换信息的过程

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 AS_1 的 BGP 发言人: “要到达网络 N_1 、 N_2 、 N_3 和 N_4 可经过 AS_2 。”



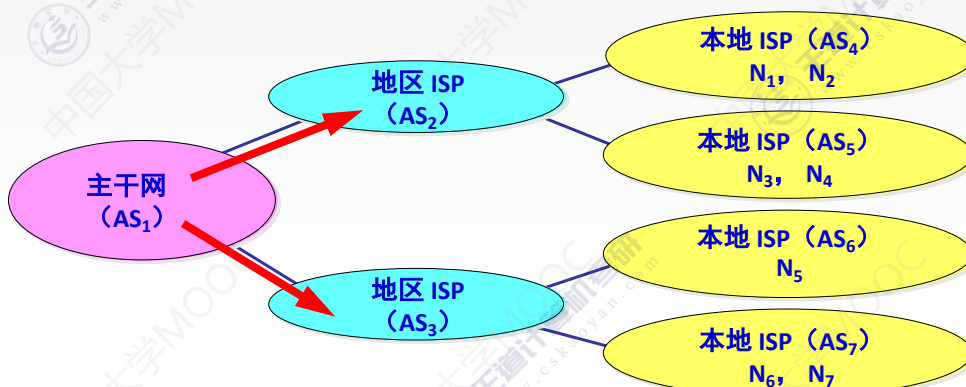
王道考研/CSKAOYAN.COM

31

BGP协议交换信息的过程

BGP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)。”

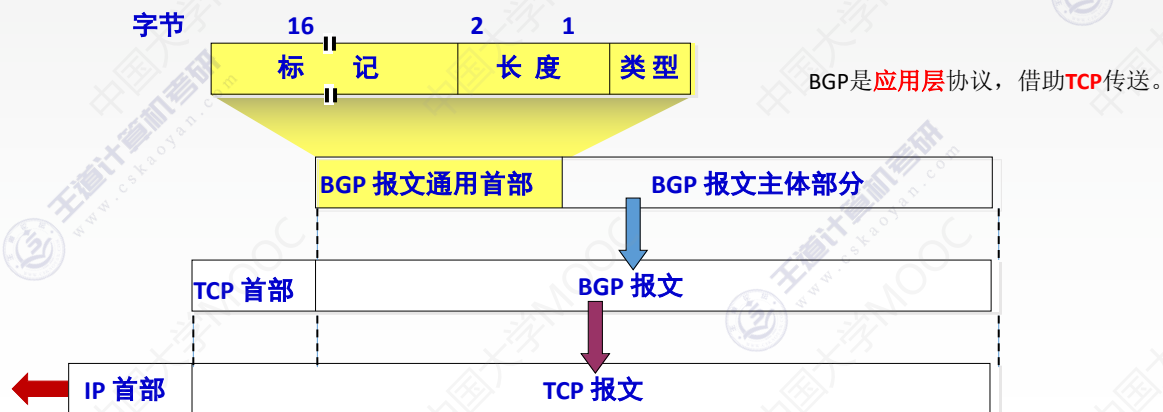


王道考研/CSKAOYAN.COM

32

BGP协议报文格式

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，即通过 TCP 传送，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。



王道考研/CSKAOYAN.COM

33

BGP协议特点

BGP 支持 CIDR，因此 BGP 的路由表也就应当包括目的网络前缀、下一跳路由器，以及到达该目的网络所要经过的各个自治系统序列。

在 BGP 刚刚运行时，BGP 的邻站是交换整个的 BGP 路由表。但以后只需要在发生变化时更新有变化的部分。这样做对节省网络带宽和减少路由器的处理开销都有好处。

王道考研/CSKAOYAN.COM

34

BGP-4的四种报文

- 1.OPEN（打开）报文：**用来与相邻的另一个BGP发言人建立关系，并认证发送方。
- 2.UPDATE（更新）报文：**通告新路径或撤销原路径。
- 3.KEEPALIVE（保活）报文：**在无UPDATE时，周期性证实邻站的连通性；也作为OPEN的确认。
- 4.NOTIFICATION（通知）报文：**报告先前报文的差错；也被用于关闭连接。

王道考研/CSKAOYAN.COM

35

三种路由协议比较

RIP是一种分布式的基于距离向量的内部网关路由选择协议，通过广播**UDP**报文来交换路由信息。

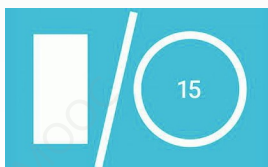
OSPF是一个内部网关协议，要交换的信息量较大，应使报文的长度尽量短，所以不使用传输层协议（如UDP或TCP），而是直接采用**IP**。

BGP是一个外部网关协议，在不同的自治系统之间交换路由信息，由于网络环境复杂，需要保证可靠传输，所以采用**TCP**。

RU



OI



BT



王道考研/CSKAOYAN.COM

36

三种路由协议比较

协议	RIP	OSPF	BGP	
类型	内部	内部	外部	
路由算法	距离-向量	链路状态	路径-向量	
传递协议	UDP	IP	TCP	
路径选择	跳数最少	代价最低	较好，非最佳	
交换结点	和本结点相邻的路由器	网络中的所有路由器	和本结点相邻的路由器	
交换内容	当前本路由器知道的全部信息，即自己的路由表	与本路由器相邻的所有路由器的链路状态	首次	整个路由表
			非首次	有变化的部分

王道考研/CSKAOYAN.COM

37

本节内容

IP组播

王道考研/CSKAOYAN.COM

38

IP数据报的三种传输方式

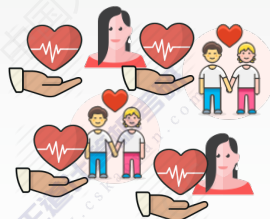
单播

单播用于发送数据包到单个目的地，且每发送一份单播报文都使用一个单播IP地址作为目的地址。是一种**点对点**传输方式。



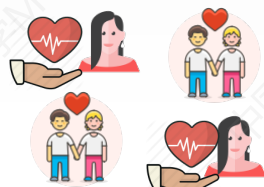
广播

广播是指发送数据包到同一广播域或子网内的所有设备的一种数据传输方式，是一种**点对多点**传输方式。



组播（多播）

当网络中的某些用户需要特定数据时，组播数据发送者仅发送一次数据，借助组播路由协议为组播数据包建立组播分发树，被传递的数据到达距离用户端尽可能近的节点后才开始复制和分发，是一种**点对多点**传输方式。

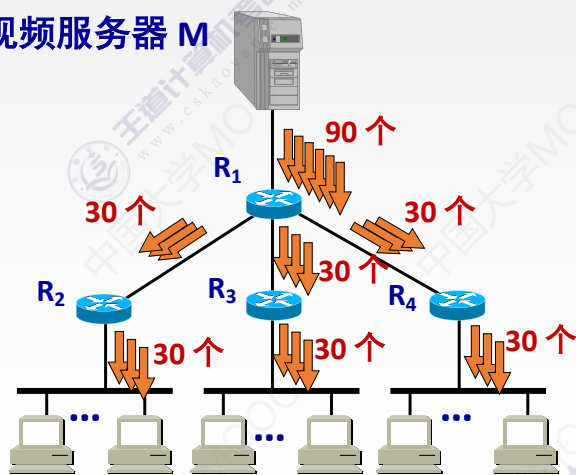


王道考研/CSKAOYAN.COM

39

IP数据报的三种传输方式

视频服务器 M



共有 90 个主机接收视频节目

单播

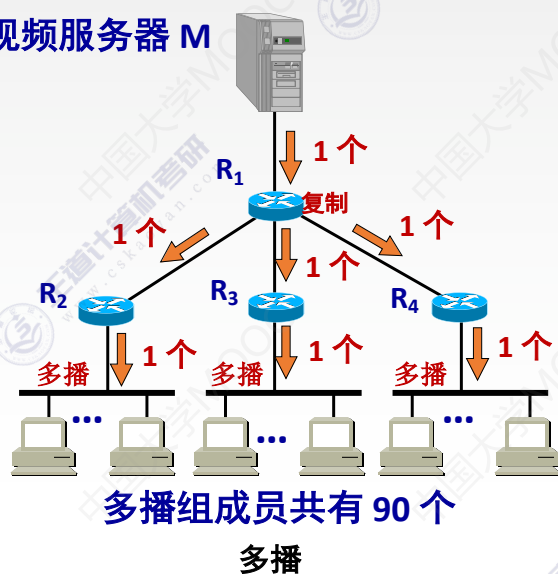
在发送者和每一接收者之间需要**单独**的数据信道。

王道考研/CSKAOYAN.COM

40

IP数据报的三种传输方式

视频服务器 M



组播提高了数据传送效率。减少了主干网出现拥塞的可能性。组播组中的主机可以是在同一个物理网络，也可以来自不同的物理网络（如果有组播路由器的支持）。

运行组播协议的路由器

王道考研/CSKAOYAN.COM

41

IP组播地址

IP组播地址让源设备能够将分组发送给一组设备。属于多播组的设备将被分配一个组播组IP地址（一群共同需求主机的相同标识）。

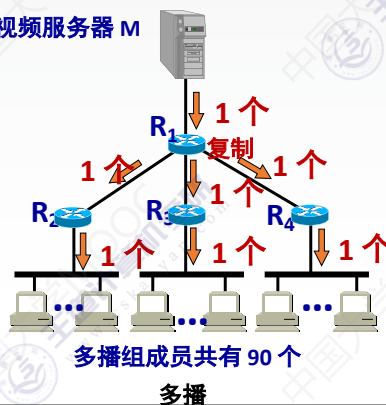
组播地址范围为224.0.0.0~239.255.255.255（D类地址），一个D类地址表示一个组播组。只能用作分组的目标地址。源地址总是为单播地址。

- 1.组播数据报也是“尽最大努力交付”，不提供可靠交付，应用于UDP。
- 2.对组播数据报不产生ICMP差错报文。
- 3.并非所有D类地址都可以作为组播地址。

视频服务器 M

因特网范围内组播

硬件组播



王道考研/CSKAOYAN.COM

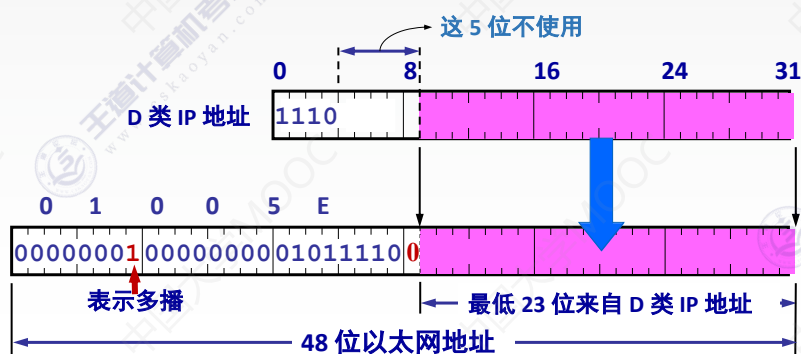
42

硬件组播

同单播地址一样，组播IP地址也需要相应的组播MAC地址在本地网络中实际传送帧。组播MAC地址以十六进制值01-00-5E打头，余下的6个十六进制位是根据IP组播组地址的最后23位转换得到的。

TCP/IP 协议使用的以太网多播地址的范围是：

从01-00-5E-00-00-00到01-00-5E-7F-FF-FF。



收到多播数据报的主机，还要在 IP 层利用软件进行过滤，把不是本主机要接收的数据报丢弃。

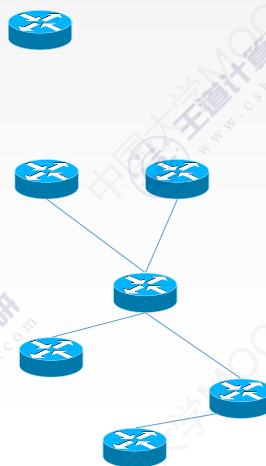
王道考研/CSKAOYAN.COM

43

IGMP协议与组播路由选择协议

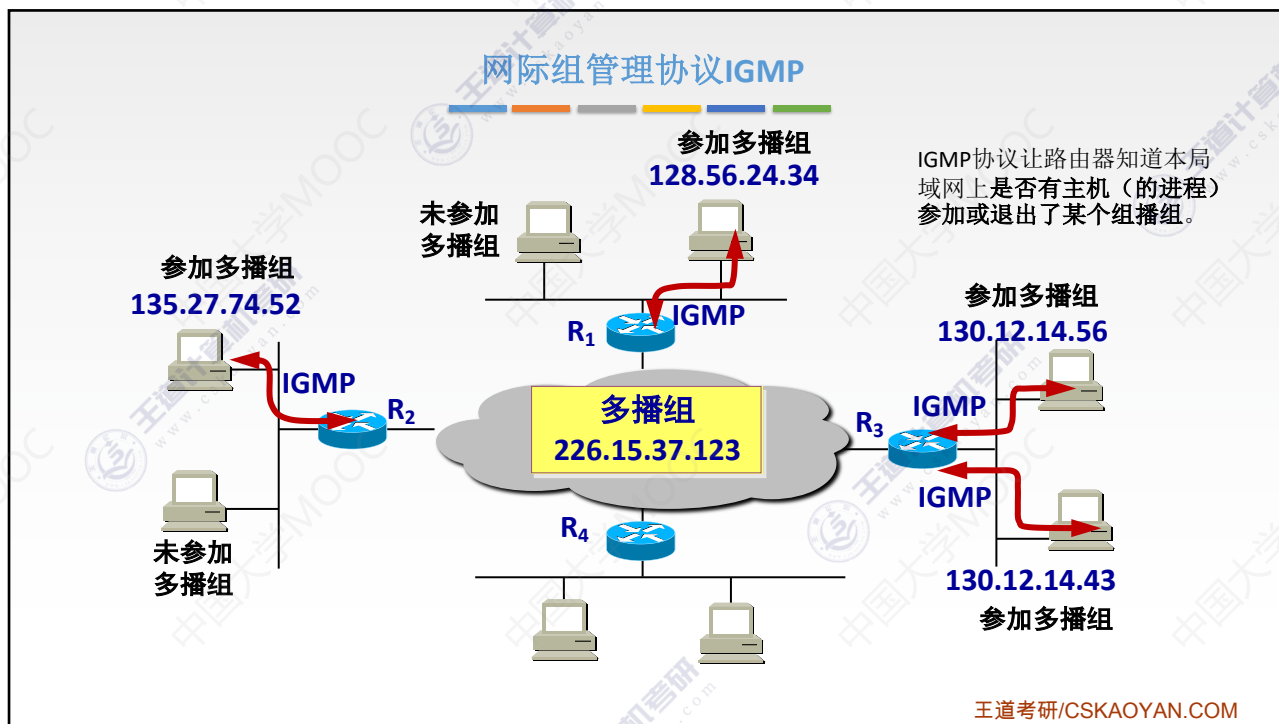
IGMP协议

组播路由选择协议

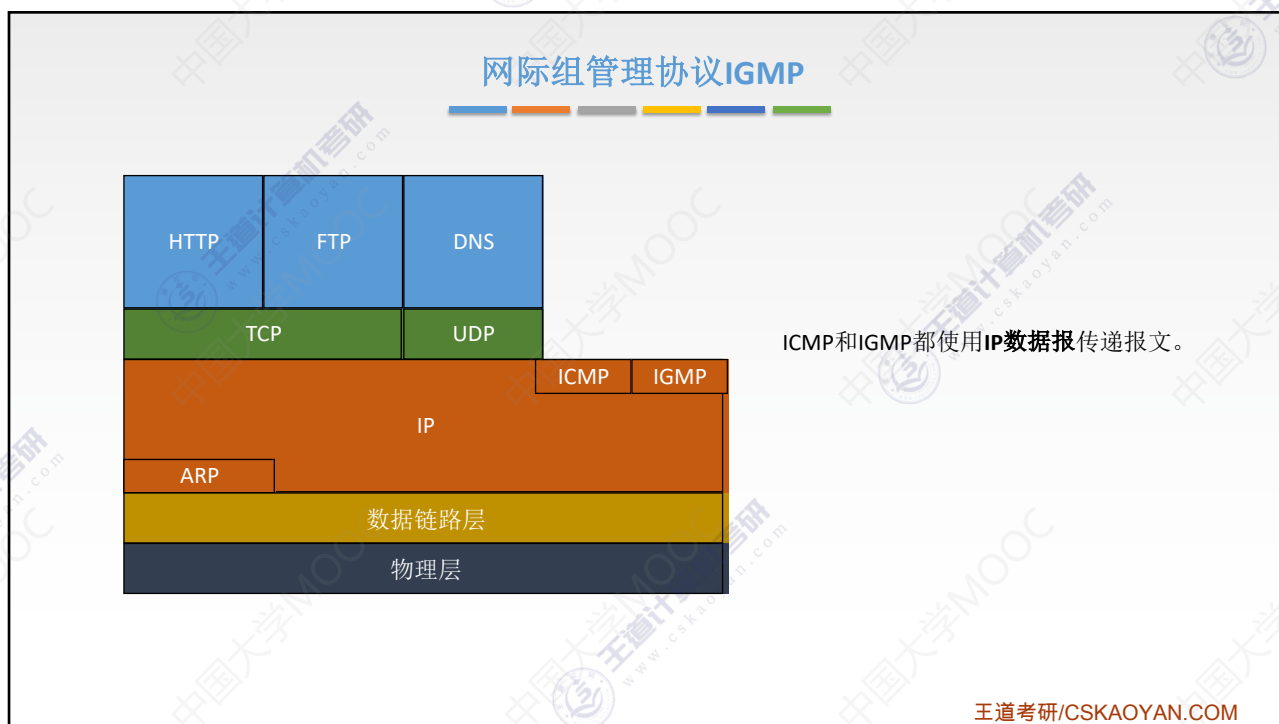


王道考研/CSKAOYAN.COM

44



45



46

IGMP工作的两个阶段

ROUND 1:

某主机要加入组播组时，该主机向组播组的组播地址发送一个IGMP报文，声明自己要称为该组的成员。
本地组播路由器收到IGMP报文后，要利用组播路由选择协议把这组成员关系发给因特网上的其他组播路由器。

ROUND 2:

本地组播路由器周期性探询本地局域网上的主机，以便知道这些主机是否还是组播组的成员。
只要有一个主机对某个组响应，那么组播路由器就认为这个组是活跃的；如果经过几次探询后没有一个主机响应，组播路由器就认为本网络上的没有此组播组的主机，因此就不再把这组的成员关系发给其他的组播路由器。

组播路由器知道的成员关系只是所连接的局域网中有无组播组的成员。

王道考研/CSKAOYAN.COM

47

组播路由选择协议



组播路由选择协议目的是找出以源主机为根节点的组播转发树。

王道考研/CSKAOYAN.COM

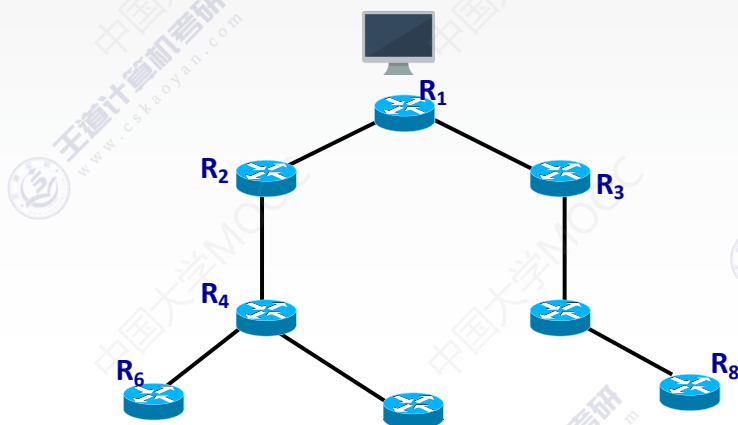
48

组播路由选择协议

组播路由协议目的是找出以源主机为根节点的**组播转发树**。

构造树可以避免在路由器之间兜圈子。

对不同的多播组对应于不同的多播转发树；同一个多播组，对不同的源点也会有不同的多播转发树。



王道考研/CSKAOYAN.COM

49

组播路由选择协议

组播路由选择协议常使用的三种算法：

基于链路状态的路由选择

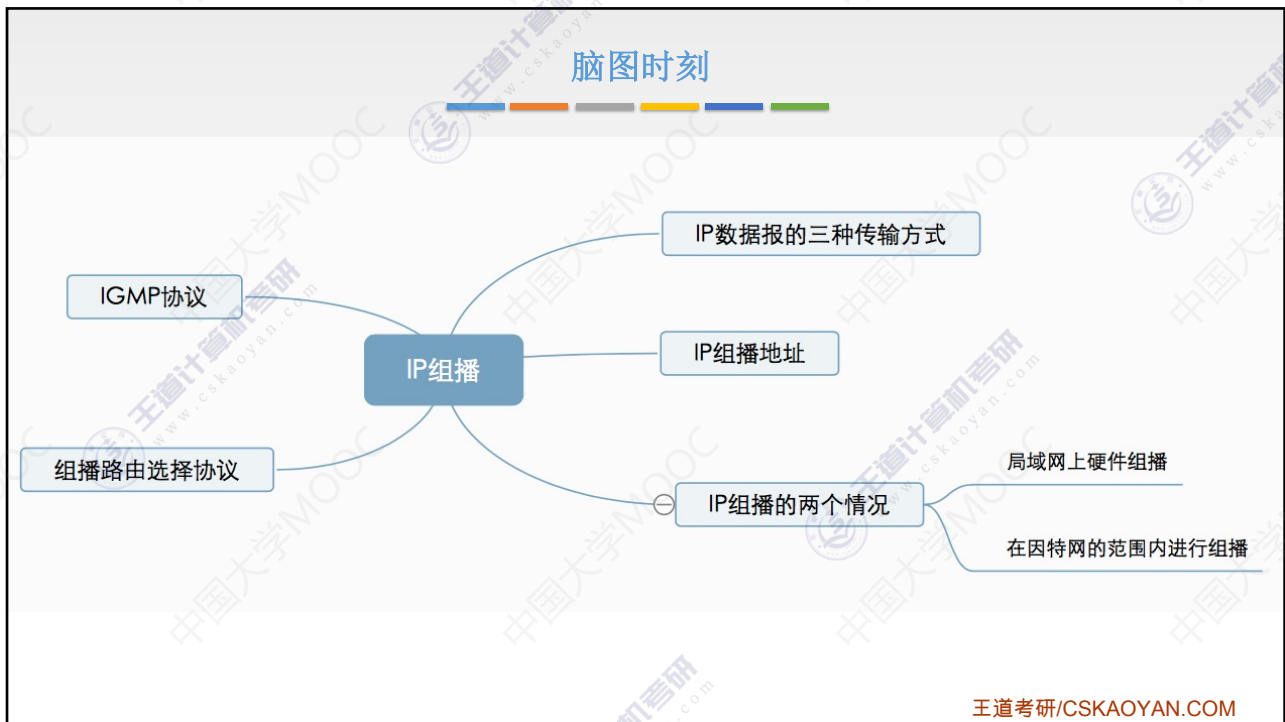
基于距离-向量的路由选择

协议无关的组播（稀疏/密集）

王道考研/CSKAOYAN.COM

50

脑图时刻



51

本节内容

移动IP

王道考研/CSKAOYAN.COM

52

移动IP



王道考研/CSKAOYAN.COM

53

移动IP

相关概念 通信过程

王道考研/CSKAOYAN.COM

54

移动IP相关概念

移动IP技术是移动结点(计算机/服务器等)以**固定的网络IP地址**，实现跨越不同网段的**漫游**功能，并保证了基于网络IP的网络权限在漫游过程中不发生改变。

移动结点 具有永久IP地址的移动设备。

归属代理（本地代理） 一个移动结点的永久“居所”称为**归属网络**，在归属网络中代表移动节点执行移动管理功能的实体叫做归属代理。

永久地址（归属地址/主地址） 移动站点在归属网络中的原始地址。

外部代理（外地代理） 在**外部网络**中帮助移动节点完成移动管理功能的实体称为外部代理。

转交地址（辅地址） 可以是外部代理的地址或动态配置的一个地址。

王道考研/CSKAOYAN.COM

55

移动IP

归属代理：村子



爷爷你们在干嘛



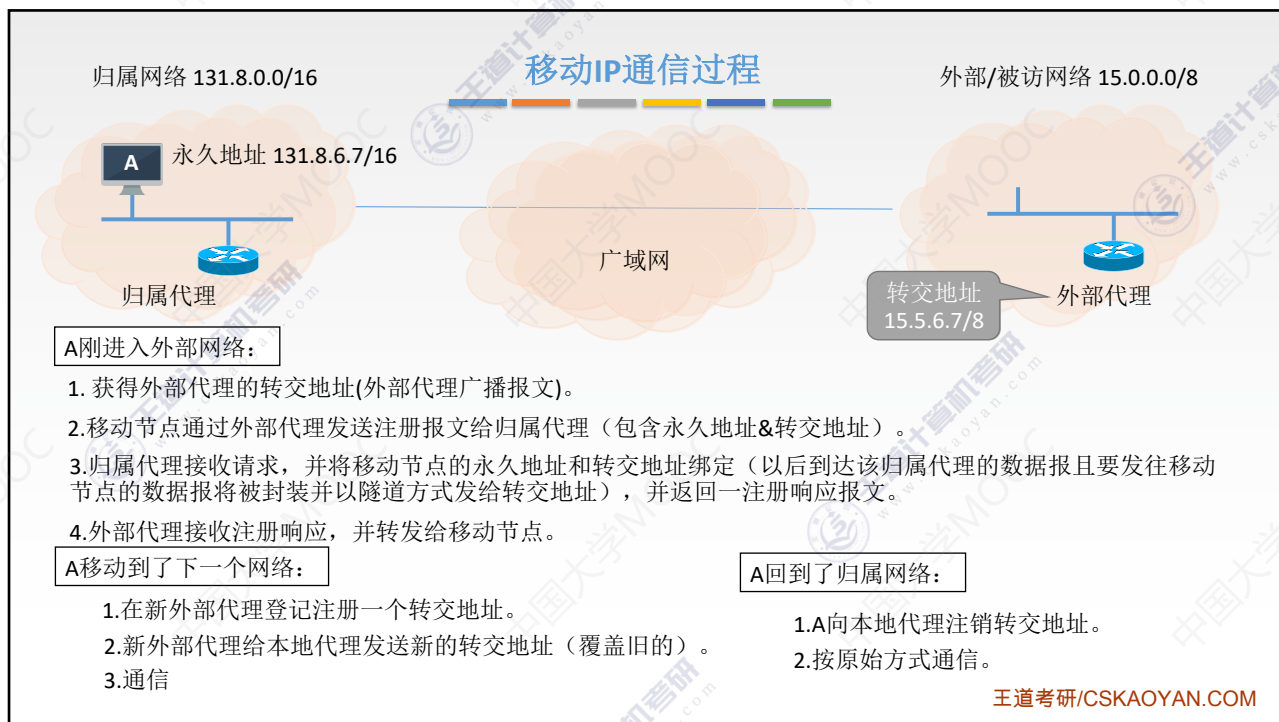
移动节点：爷爷
永久地址：爷爷的id

外部代理：盘丝洞
转交地址：盘丝洞id



王道考研/CSKAOYAN.COM

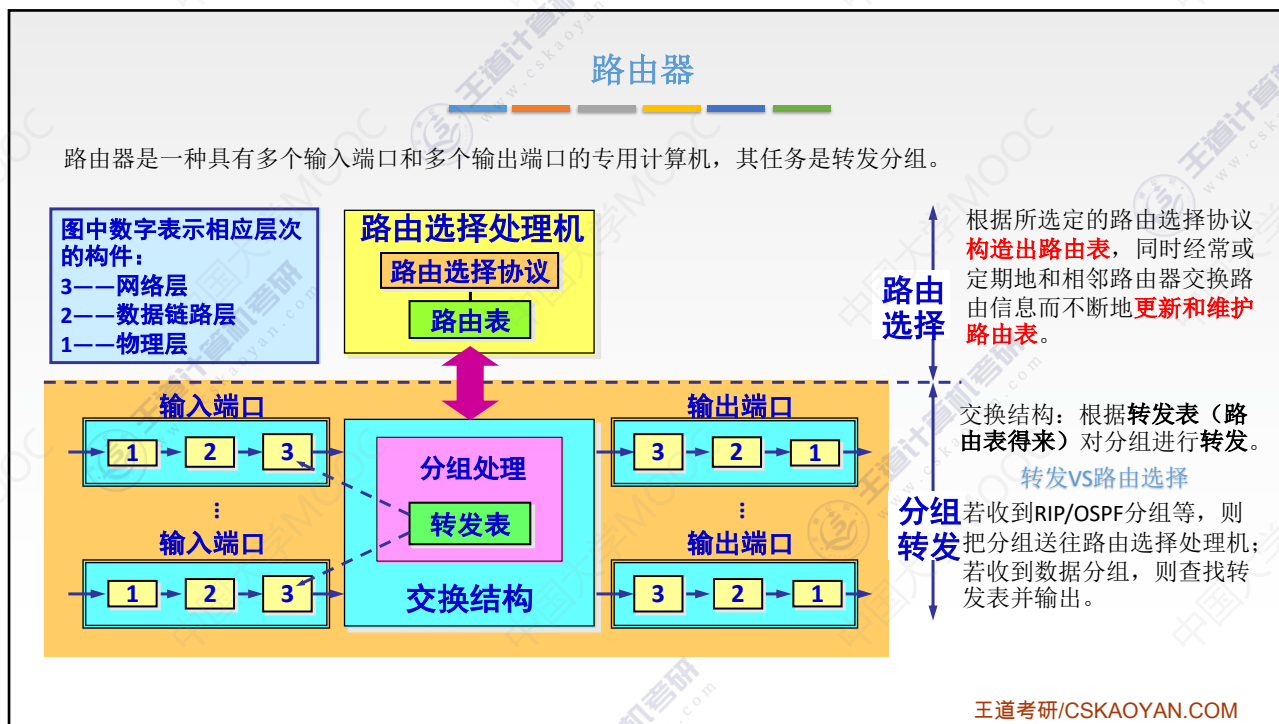
56



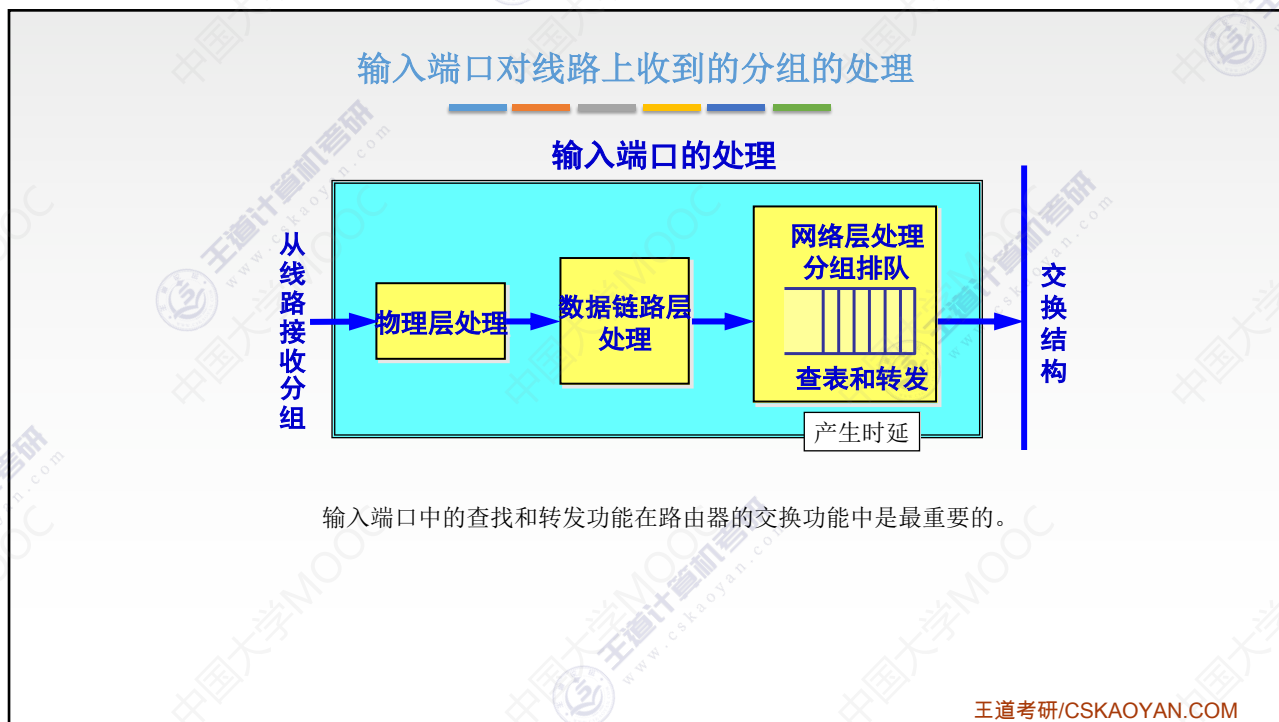
57



58

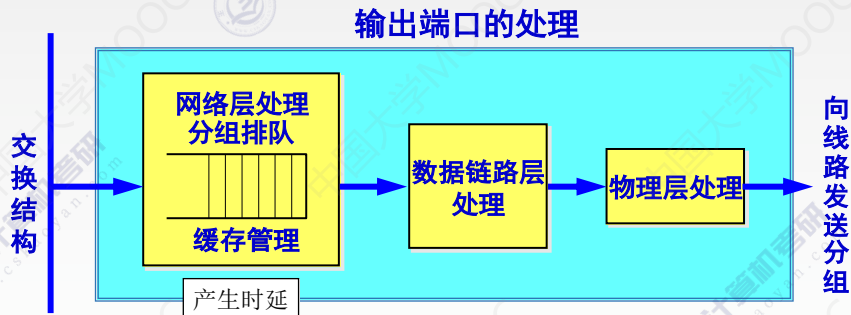


59



60

输出端口将交换结构传送来的分组发送到线路



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。

王道考研/CSKAOYAN.COM

61

三层设备的区别

路由器 可以互联两个不同网络层协议的网段。

网桥 可以互联两个物理层和链路层不同的网段。

集线器 不能互联两个物理层不同的网段。

	能否隔离冲突域	能否隔离广播域
物理层设备【傻瓜】 (中继器、集线器)	×	×
链路层设备【路人】 (网桥、交换机)	√	×
网络层设备【大佬】 (路由器)	√	√

王道考研/CSKAOYAN.COM

62

路由表与路由转发

路由表根据**路由选择算法**得出的，主要用途是路由选择，总用软件来实现。

路由表

	目的网络IP地址	子网掩码	下一跳IP地址	接口
默认路由	0.0.0.0	0.0.0.0		



转发表

转发表由**路由表**得来，可以用软件实现，也可以用特殊的硬件来实现。转发表必须包含完成转发功能所必需的信息，在转发表的每一行必须包含从要到达的目的网络到输出端口和某些MAC地址信息的映射。

王道考研/CSKAOYAN.COM