

1. 路由选择协议

1.1 静态路由选择

- 由人工配置的网络路由、默认路由、特定主机路由、黑洞路由等都属于静态路由
- 这种人工配置的方式简单、开销小；但不能及时适应网络状态（流量、拓扑等）变化
- 一般只在小规模网络中采用

1.2 动态路由选择

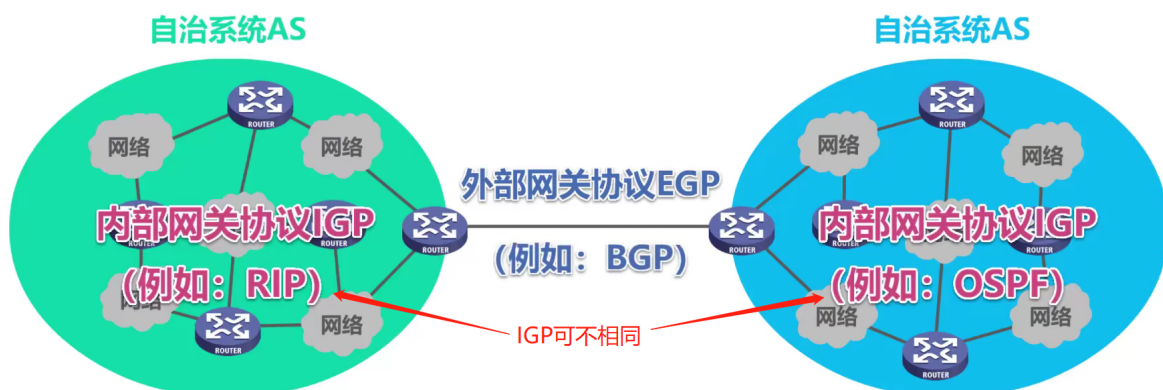
- 路由器通过路由选择协议自动获取路由信息
- 比较复杂、开销比较大；能较好地适应网络状态的变化
- 适用于大规模网络

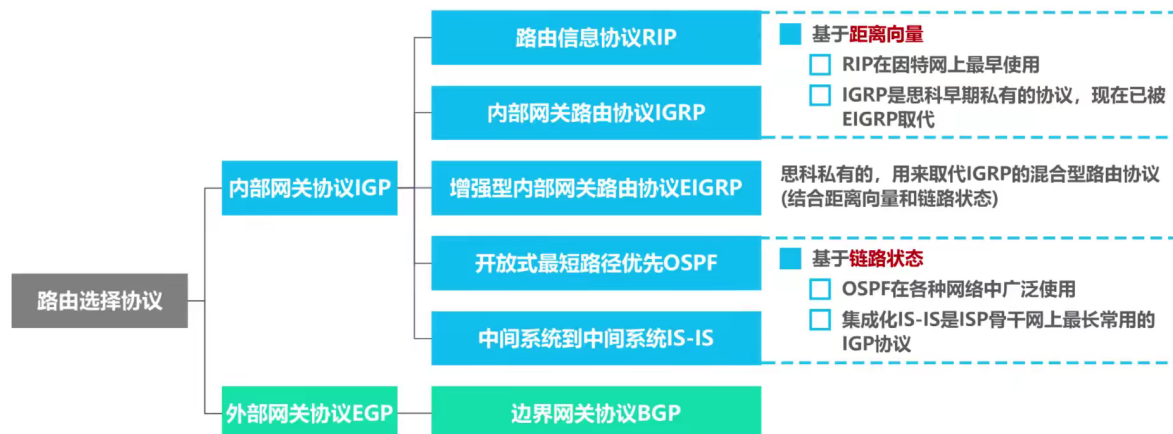
1.3 因特网所采用的路由选择协议的主要特点

- 自适应：
 - 因特网采用动态路由选择，能较好地适应网络状态的变化
- 分布式：
 - 因特网间的各路由器通过相互间的信息交互共同完成路由信息的获取和更新
- 分层次：
 - 将整个因特网划分为许多较小的自治系统AS
 - 一个较大的因特网服务提供商ISP，就可划分为一个AS，在AS内部和AS外部采用不同类别的路由选择协议分别进行路由选择

1.4 内部网关协议与外部网关协议

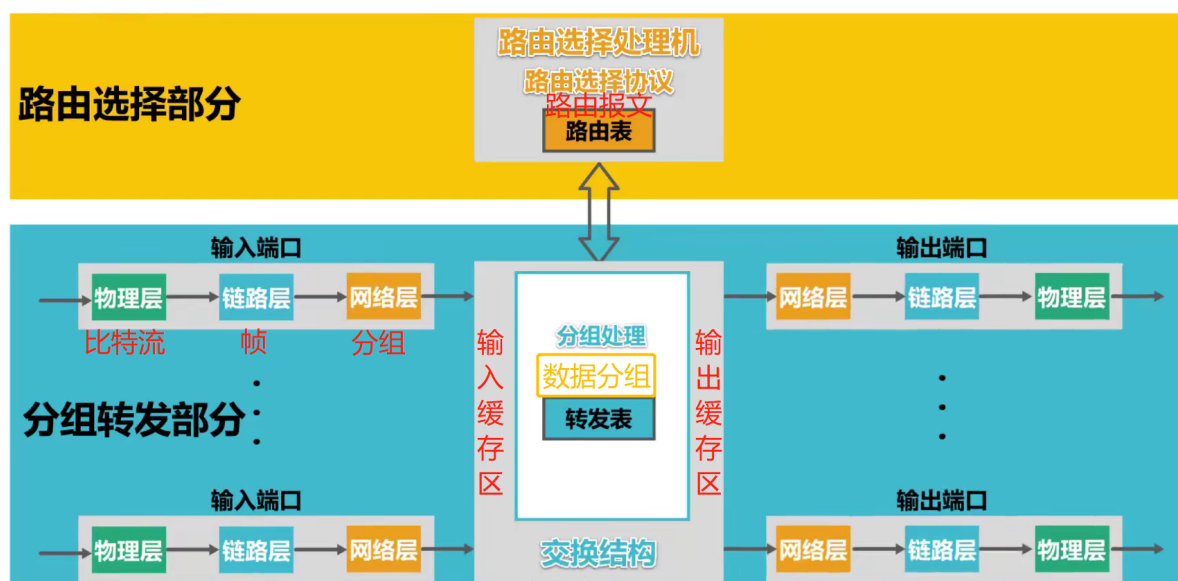
- AS间的路由选择简称为域间路由选择，采用外部网关协议EGP这个类别（不特指某具体）的路由选择协议
- AS内的路由选择简称为域内路由选择，采用内部网关协议IGP这个类别（不特指某具体）的路由选择协议
 - “网关”：因为因特网早期的RFC文档中，没有使用“路由器”，而采用“网关”这个名词。可改称ERP，IRP。





1.5 路由器的基本结构

- 路由器是一种具有多个输入端口，多个输出端口的专用计算机，任务为转发分组
- 结构可划分为两个部分：
 - 路由选择部分
 - 核心构件是**路由选择处理机**，其任务是根据所使用的路由选择协议，周期性的与其他路由器进行路由信息的交互，更新路由表
 - 处理收到的路由报文
 - 周期性的给其他路由器发送自己所知道的路由信息
 - 分组转发部分
 - 一组输入端口，交换结构，一组输出结构

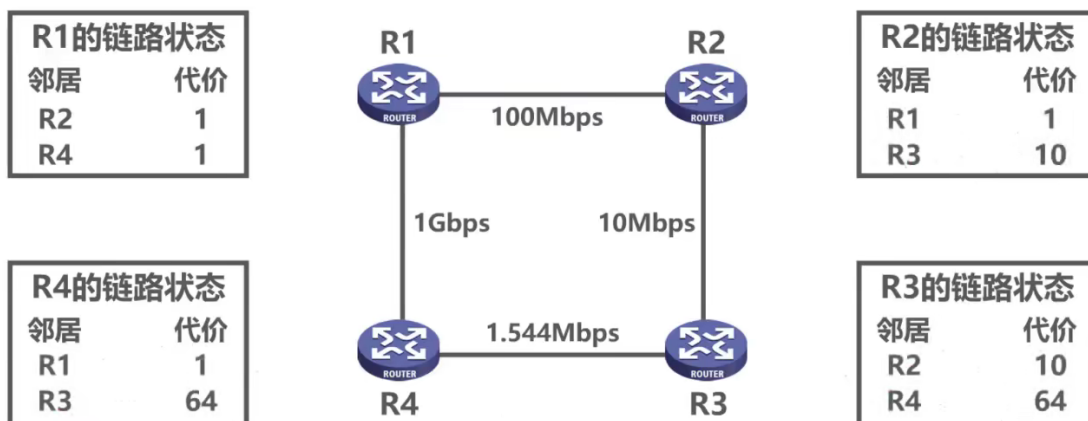


- 若为**数据分组**，根据分组首部中的目的地址进行查表（转发表）转发
 - 找不到匹配的转发条目，则丢弃该分组
 - 按匹配条目中所指示的端口进行转发
 - 输出端口网络层更新数据分组首部中某些字段的值，如将数据分组的生存时间-1，然后送交数据链路层进行封装
- 若为路由器间交互路由信息的**路由报文**，则把这种分组送交路由选择处理机
 - 路由选择处理机根据分组内容来更新自己的路由表
 - 路由表一般仅包含从目的网络到下一跳的映射
 - 路由表需要对网络拓扑变化的计算最优化
 - 转发表由路由表得出，转发表的结构应当使查找过程最优化

- 路由器的各个端口还有输入缓冲区和输出缓冲区
 - 输入缓冲区**用来暂存新进入路由器但还来不及处理的分组
 - 输出缓冲区**用来暂存已经处理完毕但还来不及发送的分组

1.6 OSPF基本工作原理

- 开放最短路径优先，是为了克服RIP的缺点在1989年开发出来的
 - 开放**表明OSPF不受某家厂商控制，公开发表
 - 最短路径优先**是指使用了Dijkstra提出的最短路径算法SPF，从算法上保证了不会产生路由环路
- OSPF不限制网络规模，更新效率高，收敛速度快
- OSPF是基于**链路状态**的，而RIP是基于距离向量的
 - 链路状态**是指本路由器都和哪些路由器相邻，以及相应的链路代价（cost）
 - 代价**表示费用、距离、时延、带宽等。这些都由网络管理人员决定
- 思科路由器中OSPF计算代价的方法 $100\text{Mbps}/\text{链路带宽}$
 - 计算结果 <1 ，记为1
 - 计算结果 >1 且有小数，舍去小数

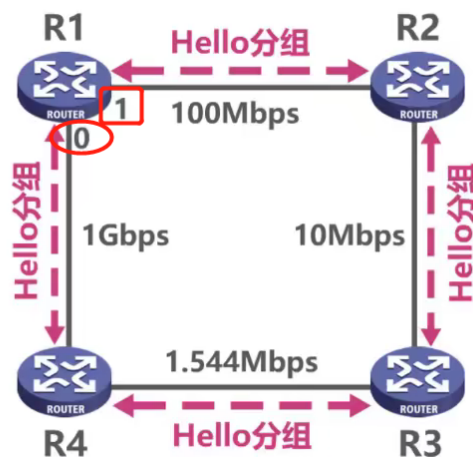


- OSPF相邻路由器之间通过交互**问候（hello）分组**，建立和维护**邻居关系**
 - Hello分组封装在IP数据报中，发往组播地址224.0.0.5
 - IP数据报中**协议号**的字段取值为89，表明数据载荷为OSPF分组
 - hello**发送周期**为10秒，40秒仍未收到来自邻居路由器的hello分组，则认为该邻居路由不可达

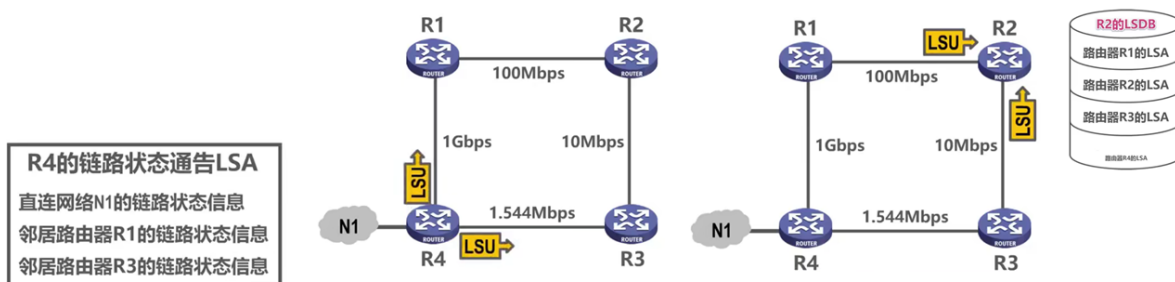
源IP 路由器接口IP	目的IP 224.0.0.5	协议号 89	OSPF 首部	OSPF分组载荷
----------------	-------------------	-----------	------------	----------

- 每个路由器都会建立**邻居表**

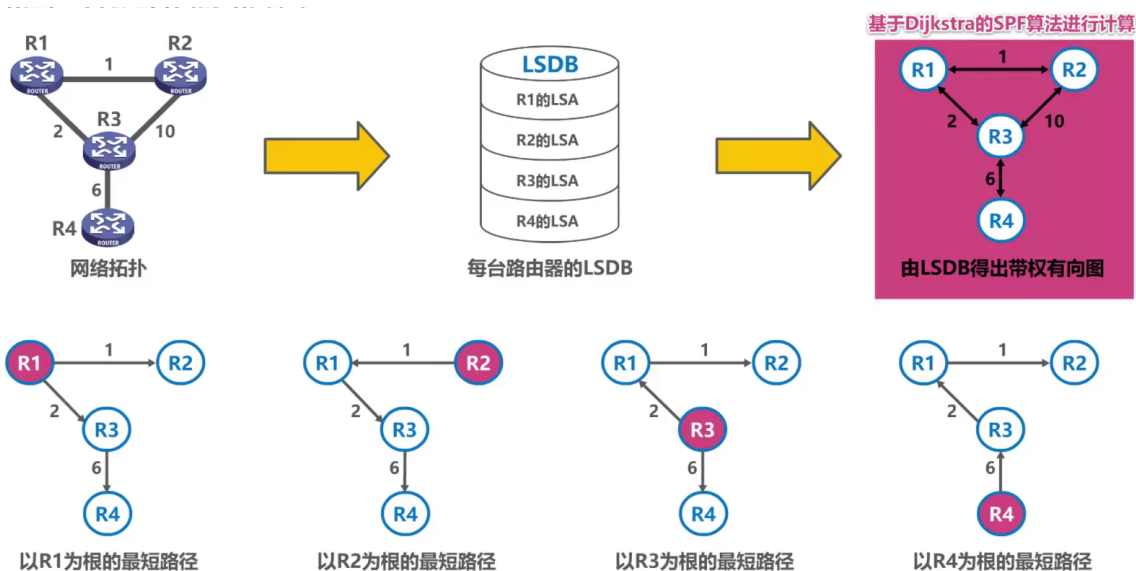
R1的邻居表		
邻居ID	接口	"死亡" 倒计时
R2	1	36秒
R4	0	18秒



- 使用OSPF的每个路由器都会产生**链路状态通告LSA** (Link State Advertisement)，LSA包含：**直连网络的链路状态信息，邻居路由器的链路状态信息**
 - LSA被封装在**链路状态更新分组LSU**中，采用**洪泛法**发送
 - 收到链路状态更新分组的路由器，将从自己其他所有接口转发该分组
 - AS中每个路由器所发送的封装有LSA的LSU会传递给系统中其他所有路由器
- 使用OSPF的每个路由器都有一个**链路状态数据库LSDB**，用于存储LSA
 - 通过各路由器洪泛发送封装有自己LSA的LSU分组，各路由器的**LSDB最终将达到一致**



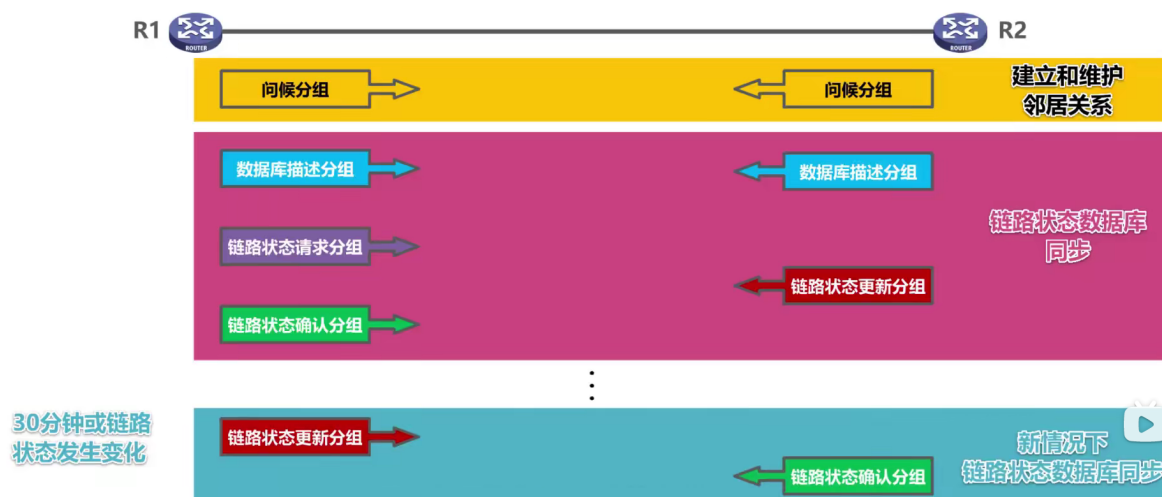
使用OSPF的各路由器**基于LSDB进行最短路径优先SPF计算**，构建出各自到达其他路由器的最短路径，即**构建各自的路由表**



2. OSPF工作过程

2.1 基本工作过程

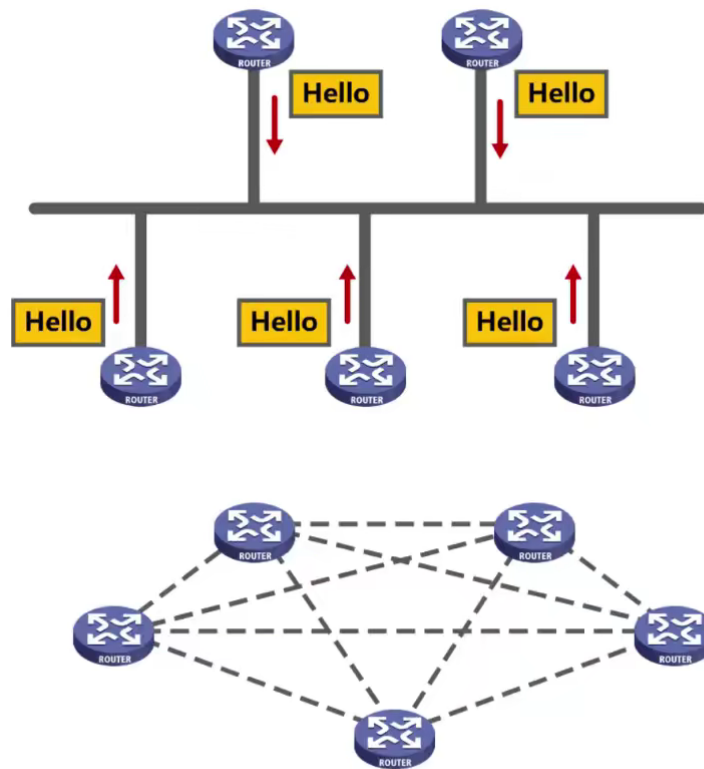
- OSPF有以下五种分组类型：
 - **问候 (hello)** 分组，用来发现和维护邻居路由器的可达性
 - **数据库描述** (Database Description) 分组，向邻居路由器给出自己的链路状态数据库LSDB中的**所有**链路状态项目的**摘要信息**
 - **链路状态请求** (Linking State Request) 分组，向邻居路由器请求发送**某些**链路状态项目的**详细信息**
 - **链路状态更新** (Link State Update) 分组，路由器使用这种分组将其链路状态进行洪泛发送，即用洪泛法对全网更新链路状态
 - **链路状态确认** (Link State Acknowledgment) 分组，是对链路状态更新分组的确认分组



- R1收到R2的数据库描述分组后，发现自己缺少其中的某些链路状态项目，向R2发送**链路状态请求分组**。
- R2收到后，将R1缺少的**链路状态项目的详细信息**封装在链路状态更新分组中发送给R1
- R1收到后，将这些信息添加到自己的LSDB中，并给R2发送**链路状态确认分组**。
- 每30分钟或链路状态发生变化时，路由器都会发送**链路状态更新分组**
- 收到该分组的其他路由器将洪泛转发该分组，并给该路由器发回**链路状态确认分组**

2.2 多点接入网络邻居关系的建立

- 若不采用其他机制，将会产生大量的多播分组，导致每个路由器要向其他 (n-1) 个路由器发送**问候分组**和**链路状态更新分组**

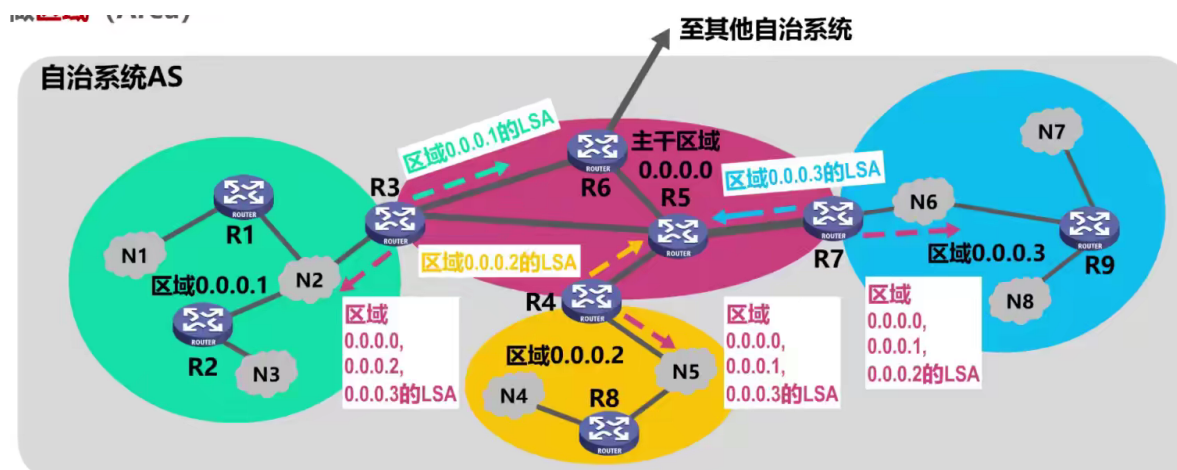


邻居关系数量: $\frac{n(n-1)}{2}$

- 采用选举指定路由器DR和备用的指定路由器BDR的方法
 - 所有非DR/BDR只与DR/BDR建立邻居关系，邻居关系数量降至 $2(n-2)+1$
 - 非DR/BDR之间不能直接交换信息
 - DR出现问题，则由BDR顶替DR
- 选举DR/BDR方法，各路由器之间交换一些选举参数如：路由器优先级、路由器ID、接口IP地址等，然后根据选举规则选出DR/BDR。（与交换机生成树协议选举根交换机类似）

2.3 大规模网络

- OSPF将AS划分为若干个更小的范围，叫做**区域 (Area)**。
 - 每个区域都有一个32bit的区域标识符，用点分十进制表示
 - 主干区域用于连通其他区域，必须为0.0.0.0
 - 其他区域标识符不能为0且互不相同
 - 每个区域规模不应太大，一般所包含的路由器<200
- 好处：把利用洪泛法交换链路状态信息的范围局限于每一个区域而非整个AS,减小了整个网络的通信量。



- 术语：

- 如果路由器的所有接口都在同一个区域内，则该路由器称为**区域内路由器IR**(Internal router)，如R1, R2, R8, R9
- **区域边界路由器ABR**(Area Border Router)：R3, R4, R7，其一个接口用于连接自身所在区域，另一个接口用于连接主干区域（可看作BBR）
- **主干路由器BBR**(Backbone router)：R3, R4, R5, R6, R7
- 自治系统边界路由器ASBR（AS Border Router）：R6

分层次划分区域的方法，虽使交换信息的种类增多、OSPF协议更加复杂，但使每个区域内部路由信息的通信量大大减小，从而使OSPF协议能够用于规模很大的AS中。