



网络层：IP组播、 移动主机及自组织网络

刘志敏

liuzm@pku.edu.cn

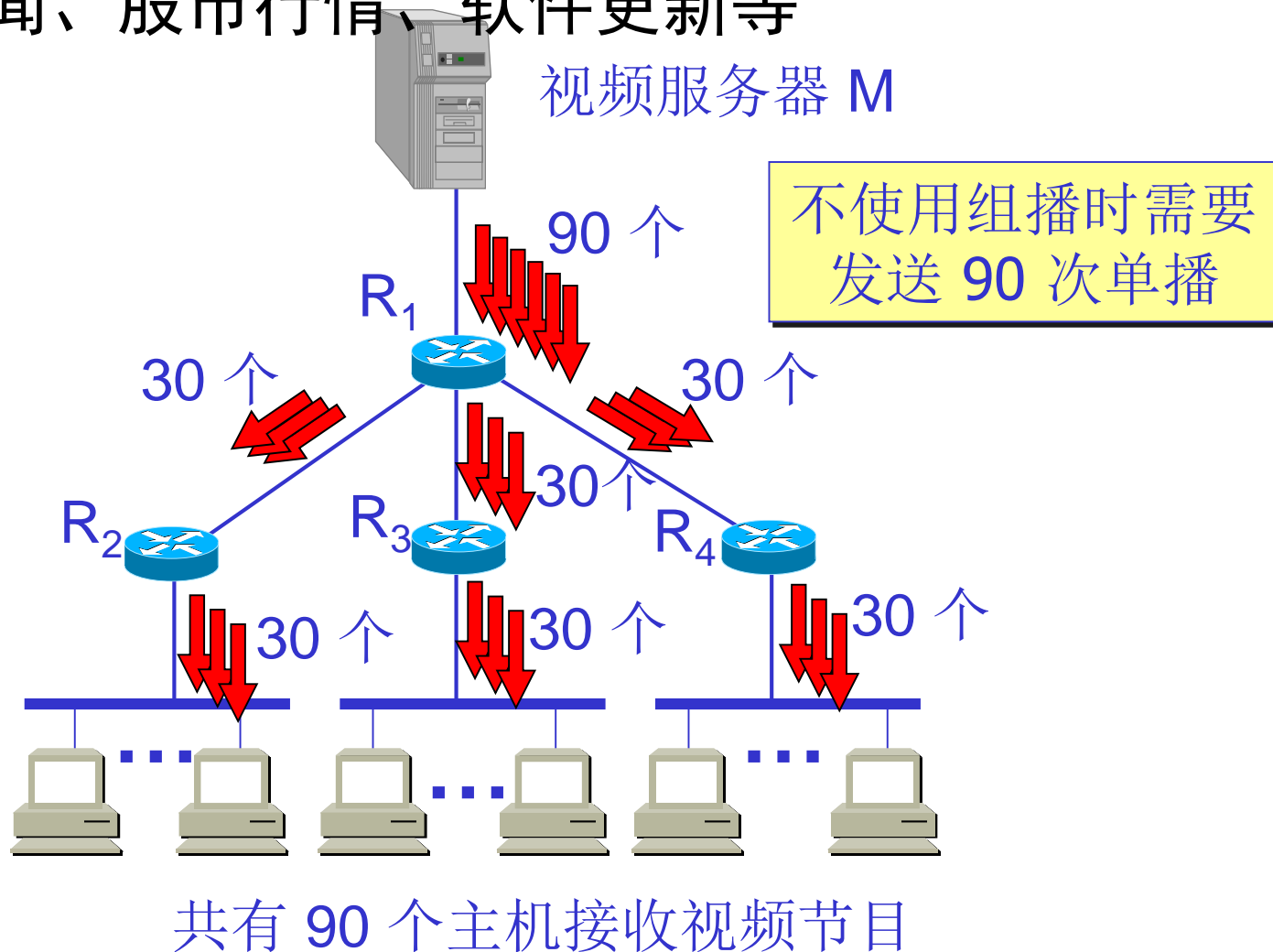


再议路由问题

- 路由与转发：
 - 路由器执行路由协议与算法、形成路由表及转发表
 - 最小距离路由、距离矢量路由
 - 路由器接收分组、查表选择转发端口
- IP组播及路由
- 移动主机路由
- 自组织网络路由

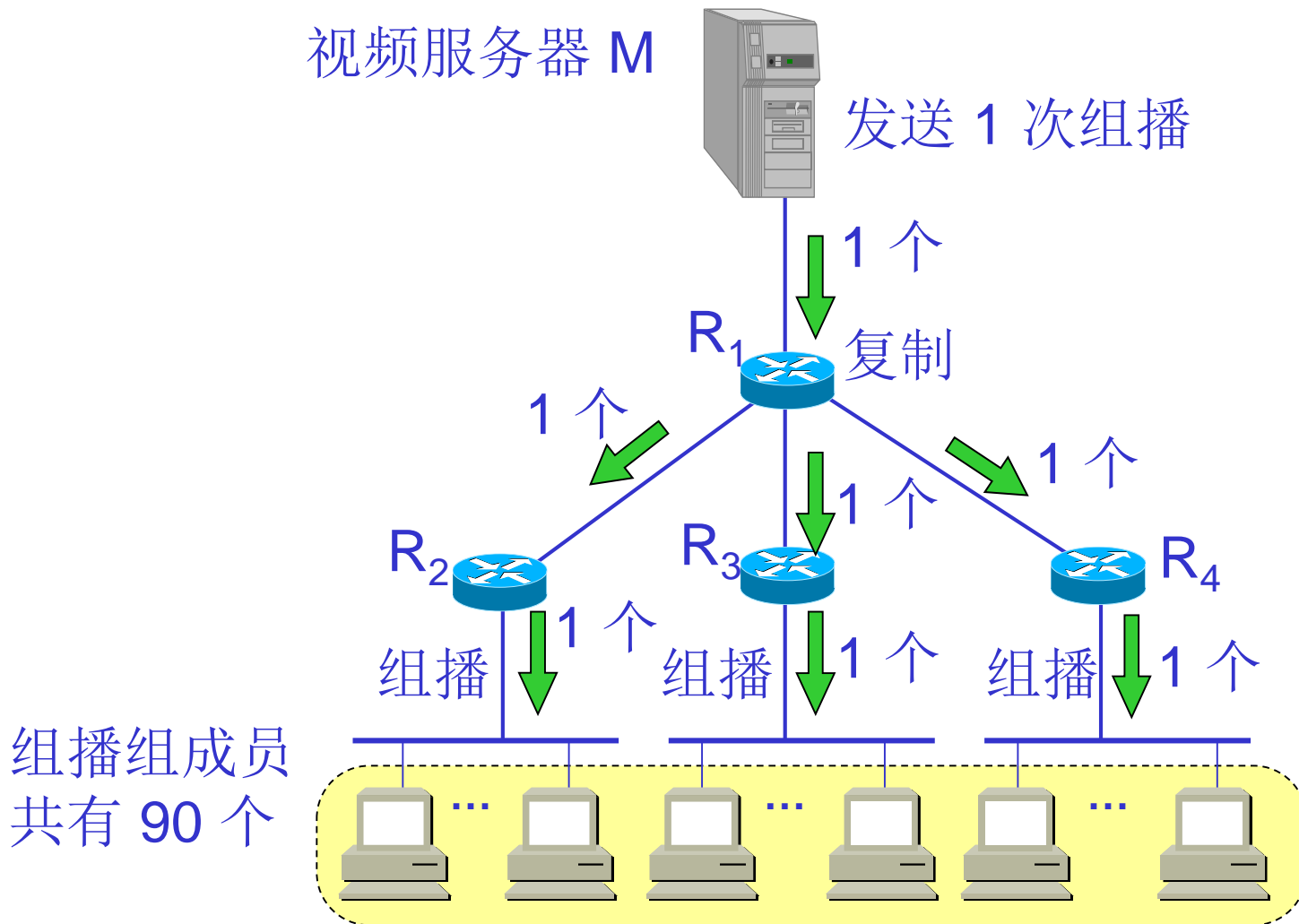
IP 组播 (Multicast)

- 一点到多点，实时信息交付，如体育赛事视频、新闻、股市行情、软件更新等



IP 组播的概念

- 减少网络中资源的消耗





IP 组播的特点

(1) 组播使用组播地址：使用IP的D类地址支持组播。

组播地址只能用于目的地址，而不能用于源地址。

(2) 永久组地址：由互联网数字分配机构 IANA

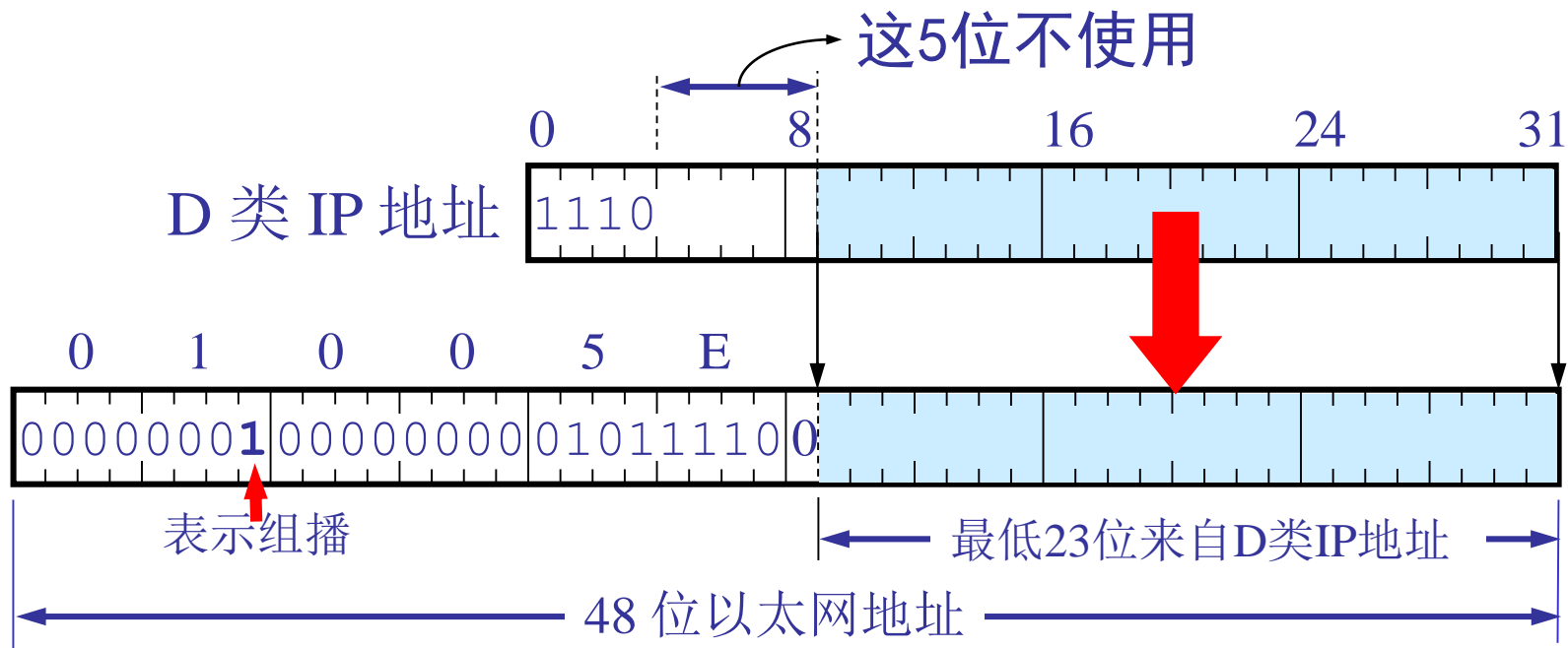
(Internet Assigned Numbers Authority) 负责分配。

(3) 组成员是动态的

(4) 使用硬件进行组播

在局域网上进行硬件组播

- 以太网硬件组播地址：首字节的最低位为1；
01-00-5E-00-00-00~01-00-5E-7F-FF-FF，
23位用于建立与D类地址的对应关系
- D类IP地址，可供分配的占28位，其中前5位不用来构成以太网硬件地址。
- 硬件接收组播帧，软件对IP组播地址进行过滤





IGMP和组播路由选择协议

1. IP组播需要两种协议

- **网际组管理协议IGMP**(Internet Group Management Protocol): 使路由器获得组播组的成员信息。
 - 一个组播组用一个D类地址标识
- **组播路由选择协议**: 使组播路由器之间协同工作, 用最小代价将组播数据报传送给所有的组成员



IGMP 分两个阶段

(1) 主机加入或离开组播组：

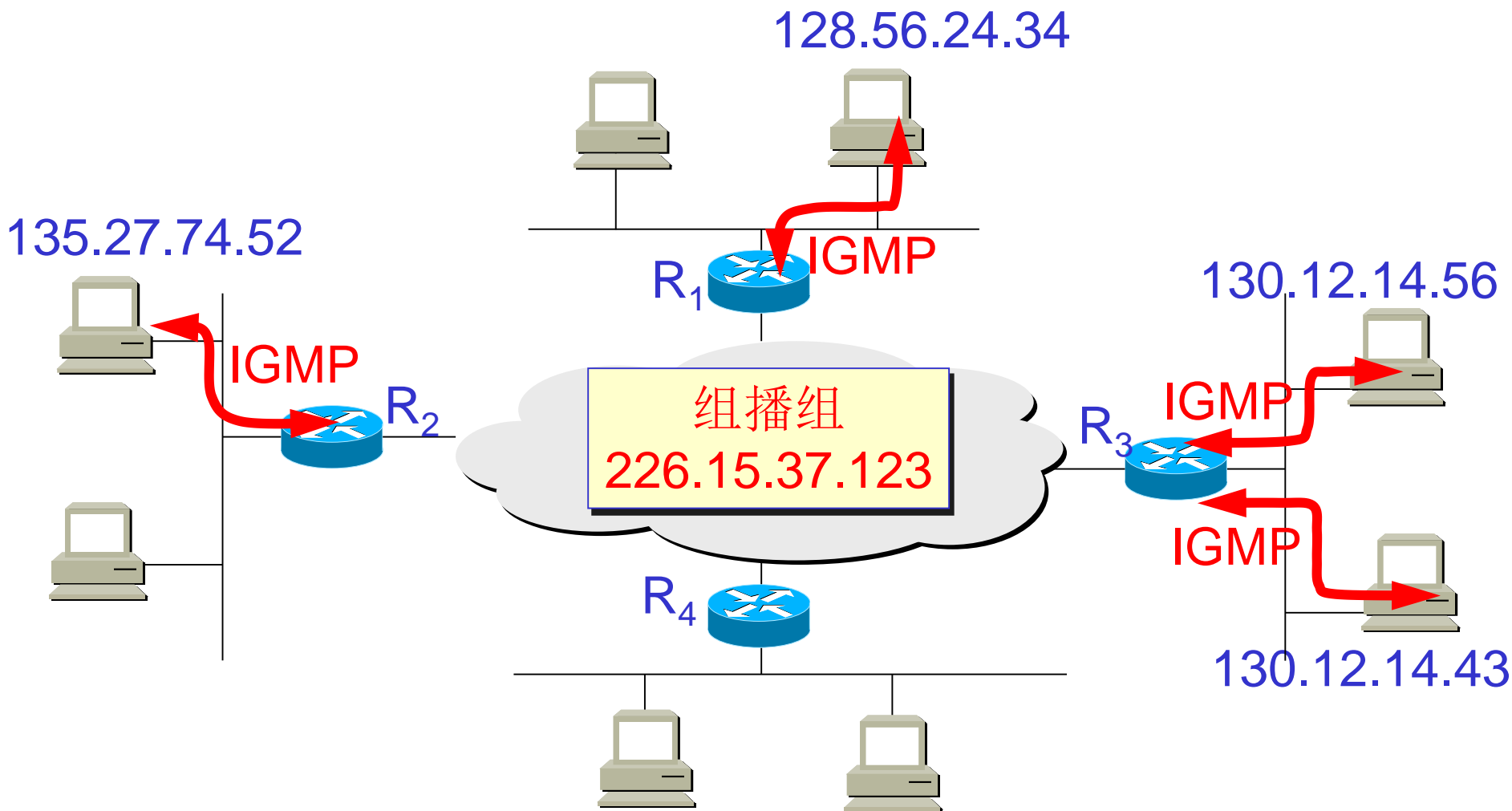
- 当主机加入组播组时，向组播组对应的某一D类地址发送IGMP成员报告报文，本地组播路由器收到IGMP报文后，将组成员关系转发给其他组播路由器。
- 当主机离开组播组时，发送IGMP成员离开报文

(2) 组播路由器维护组成员信息：周期性地发送组探测报文，因为组成员是动态的

- 某个组只要有一个主机响应，组播路由器就认为该组是活跃的；
- 若一个组在经过几次探测后该组仍然没有一个主机响应，则不再向其他组播路由器转发成员信息

IGMP使组播路由器获得组播成员信息

- 三类报文：（主机发送）报告和离开、（路由器发送）探寻





IGMP 的具体措施

- 主机和组播路由器间的通信都使用IP组播，IP的目的地址为D类地址
- 组播路由器周期性地发送询问报文（，其目的地址为组播地址），以探询组成员关系，默认间隔为125s
- 当同一个网络上有几个组播路由器时，可选择其中一个发送询问报文
- 在IGMP询问报文中有一个数值 N ，指明最长响应时间为 $N \times 0.1s$ ；
- 主机随机选择【0~ N 】时间发送成员报告
- 组内每个主机都监听成员报告，只要检测到本组其他主机已经发送了成员报告，则不再发送成员报告

IGMP: 来自https://en.wikipedia.org/wiki/Internet_Group_Management_Protocol

IGMPv2 packet structure^[8]

+	Bits 0–7	8–15	16–31
0	Type	Max Resp Time	Checksum
32	Group Address		

Type:

Membership Query (0x11),
Membership Report (IGMPv1: 0x12, IGMPv2: 0x16, IGMPv3: 0x22),
Leave Group (0x17)

Max Resp Time

Specifies the time limit for the corresponding report. The field has a resolution of 100 milliseconds. This field is meaningful only in Membership Query (0x11).

Group Address

This is the multicast address being queried when sending a Group-Specific or Group-and-Source-Specific Query. The field is zeroed when sending a General Query. The message is sent to following IP addresses:

IGMPv2 destination address^[9]

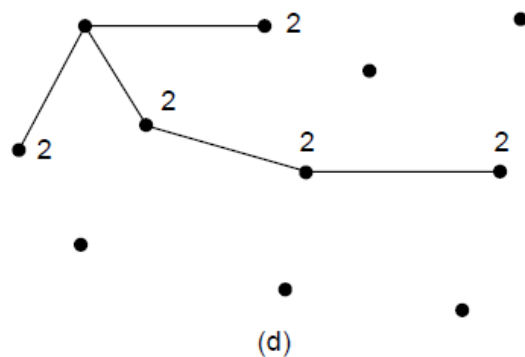
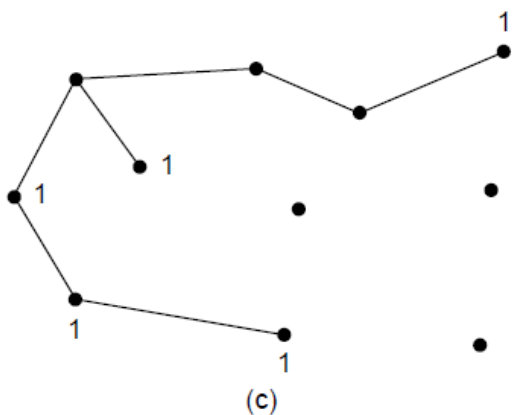
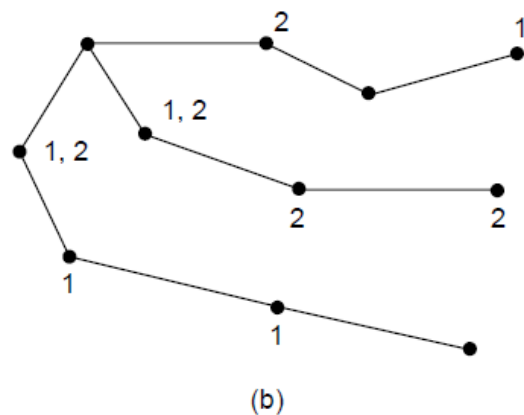
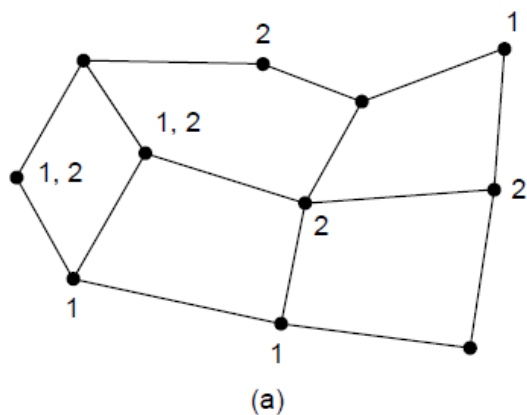
Message Type	Multicast Address
General Query	All hosts (224.0.0.1)
Group-Specific Query	The group being queried
Membership Report	The group being reported
Leave Group	All routers (224.0.0.2)



组播路由选择

- 组播组中的成员是动态变化的
- 组播路由选择是要找出以源主机为根结点的组播转发树
- 不同的组播组对应于不同的组播转发树。同一个组播组，对不同的源点也会有不同的组播转发树
- 如何转发组播数据报？
 - 基于生成树：泛洪+剪枝（修剪广播生成树）
 - 基于核心树

组播路由及其问题



(a) 网络示例：2个组播组 (b) 左侧路由器的一个生成树

(c) 组1的组播树 (d) 组2的组播树

组播组不同，组播树也不同；一个路由器需要维护多个组播树



转发组播数据报的方法

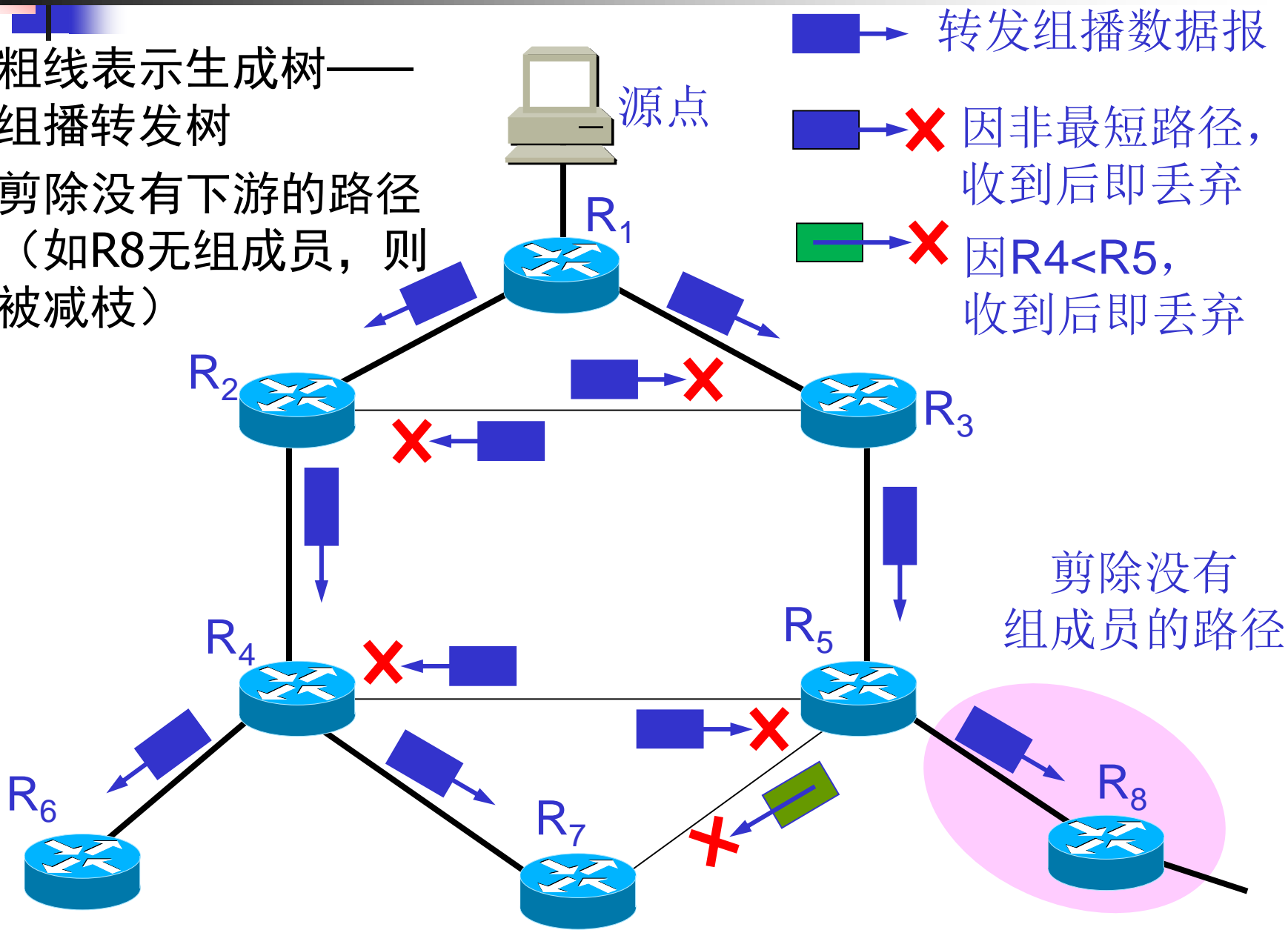
(1)基于生成树

- 路由器转发组播数据报使用泛洪（即广播）。为避免兜圈子，采用**反向路径转发RPF** (Reverse Path Forwarding)的策略
- **RPF要点：**
 - 路由器收到组播数据报时，先检查是否从源点经最短路径传送来的
 - 根据IP分组头部信息TTL？路由表中的代价？
 - 若是，就向除进入方向之外的其他方向转发；
 - 否则就丢弃；
 - 如果存在几条最短路径，则只选择IP地址最小的路由器

反向路径转发RPB和剪除

粗线表示生成树——
组播转发树

剪除没有下游的路径
(如R8无组成员, 则
被减枝)



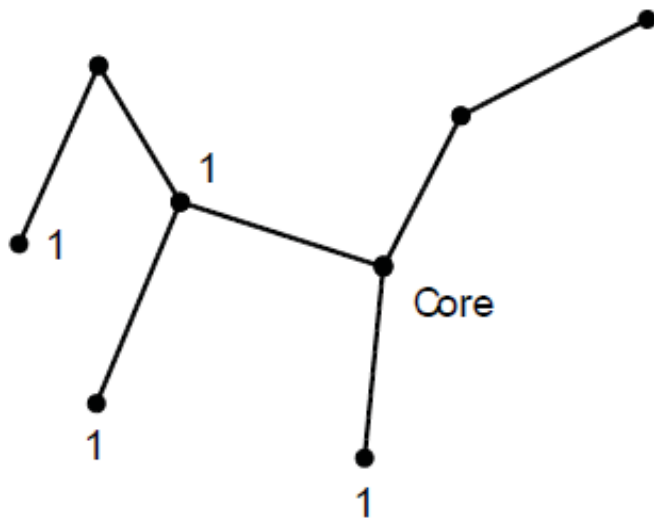
转发组播数据报的方法

- 基于核心树：计算某个组的单棵生成树，方法是全部的路由器同意某一路由器作为核心，成员发送数据分组以建立这棵树

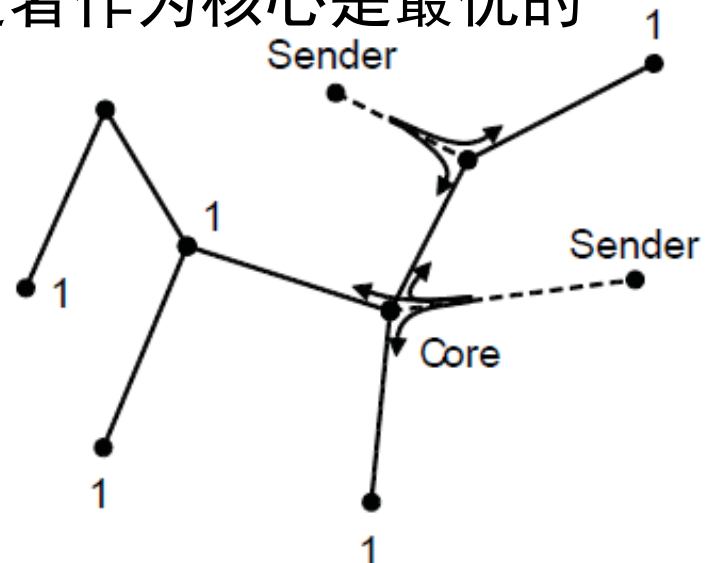
(a) 组1的核心树：成员先发送组播分组给核心

(b) 由核心发送组播分组给组1的各成员

若发送者距离核心较近，基于核心树的分发最优的；
若发送者距离核心较远，将发送者作为核心是最优的



(a)



(b)



几种组播路由选择协议

- 距离向量组播路由协议DVMRP (Distance Vector Multicast Routing Protocol)
 - 基于反向路径转发和剪除算法
- 组播MOSPF (Multicast Extensions to OSPF)
 - 基于反向路径转发和剪除算法
 - 每个路由器为每个发送者构造一棵剪除后的生成树
- 协议独立组播PIM(Protocol Independent Multicast)
 - 基于核心树，每个路由器为每个组保存一棵树
- 基于核心的转发树CBT (Core Based Tree)



组播路由选择算法比较

- 网络规模、组播的密度
- 组播密度大：基于生成树（广播+剪枝）
- 组播密度小：基于核心树
- “基于生成树”与“基于核心树”的比较
 - 若组播组总数是 n ，每个组的结点数为 m
 - 基于生成树，每个路由器维护 nm 棵生成树
 - m 个节点中，以每个节点作为根都有一个修剪的生成树； n 个组相当于有 n 个图；因此共有 nm 棵树。
 - 其优点是，可以建立一个以组播服务器（可位于网络任何位置）为根的组播树
 - 基于核心树，每个路由器维护 n 棵核心树；若发送者距离核心较近，则基于核心树的分发是最优的；否则，需要经过多跳转给核心

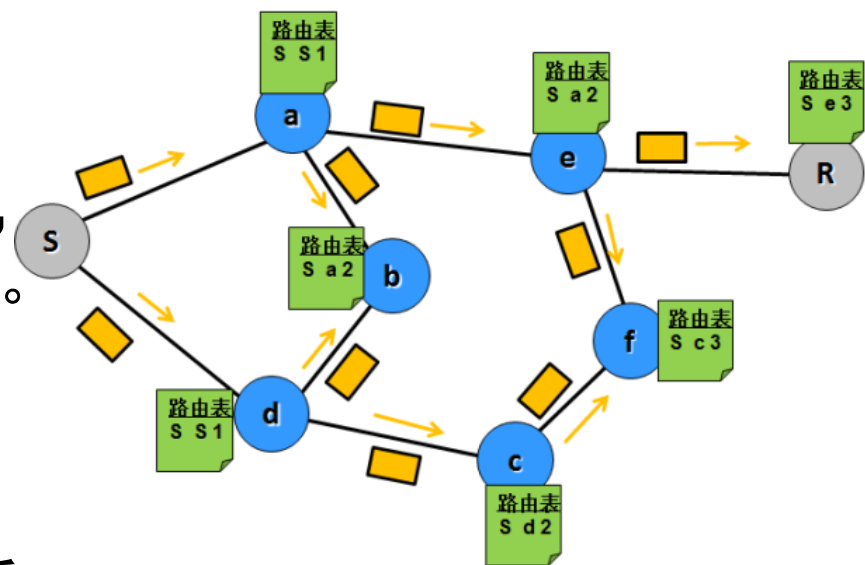
练习题

- IP组播树的构造采用一种称为反向路径转发（RPF）技术。在如图所示的网络中，假设S为组播源，R为组成员。各路由器给出的路由表部分信息为：

目的地、下一跳、距离值。

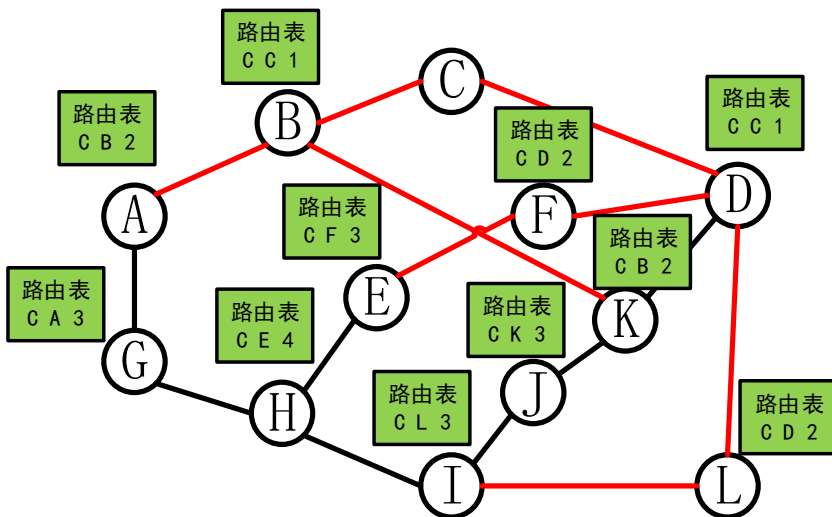
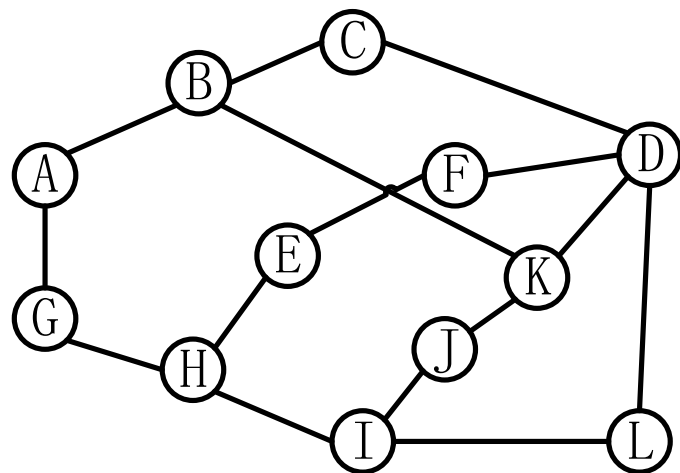
当S以泛洪方式发出一个组播

数据分组，节点b和f将收到该分组的多个拷贝。试问b和f是如何转发收到的分组？结合图例说明RPF技术。



练习题

- 在如图所示的网络中，组播成员分布在路由器A、B、C、D、E、F、I、K上，采用反向路径转发（RPF）技术。请给出路由器C的组播生成树。提示：
 - 先写出各路由器的路由表
 - 再基于RPF给出组播生成树

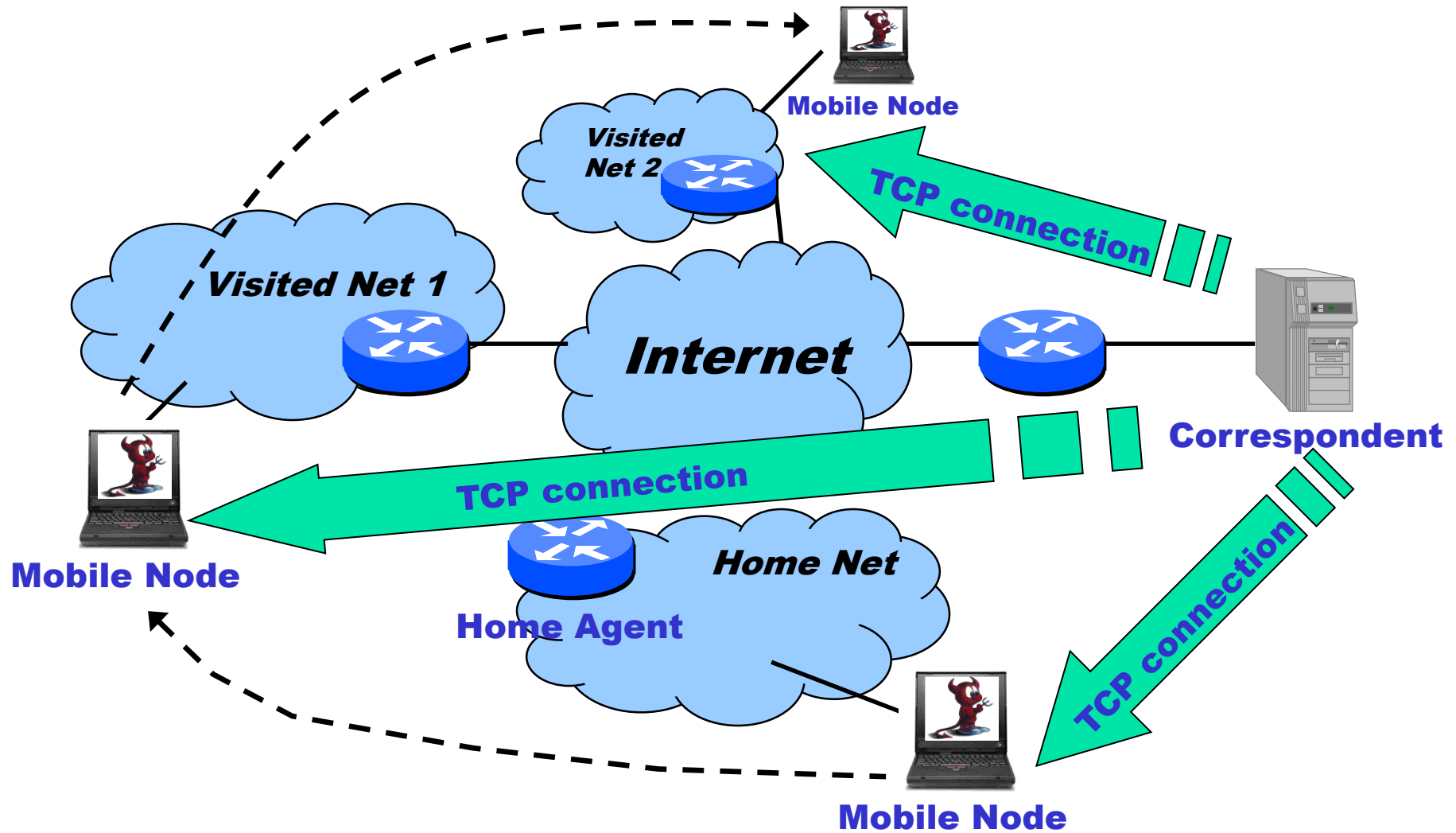




再议路由问题

- 路由与转发：
 - 路由器执行路由协议与算法、形成路由表及转发表
 - 最小距离路由、距离矢量路由
 - 路由器接收分组、查表选择转发端口
- IP组播及路由
- 移动主机路由
 - 网络中的路由器不移动，结点移动，采用移动IP
- 自组织网络路由
 - 网络中的所有节点都移动
 - MANET(Mobile AdHoc)
 - WSM((Wireless Sensor Networks)

移动主机路由





使用两个IP地址及其问题

■ IP 地址与应用相关

- IP分组交换是根据分组的目的IP地址进行路由
- IP地址既标识主机又标识主机上的应用程序
- 当改变主机与网络连接点时，根据IP路由机制，要求改变IP地址；而改变IP地址，则导致应用程序与网络之间的连接中断，因为基于UDP/TCP的进程间通信，网络进程标识使用套接字（IP，端口）
- 若采用DHCP动态IP地址分配，则在移动切换过程中，改变了IP，则出现通信中断——再连接的问题

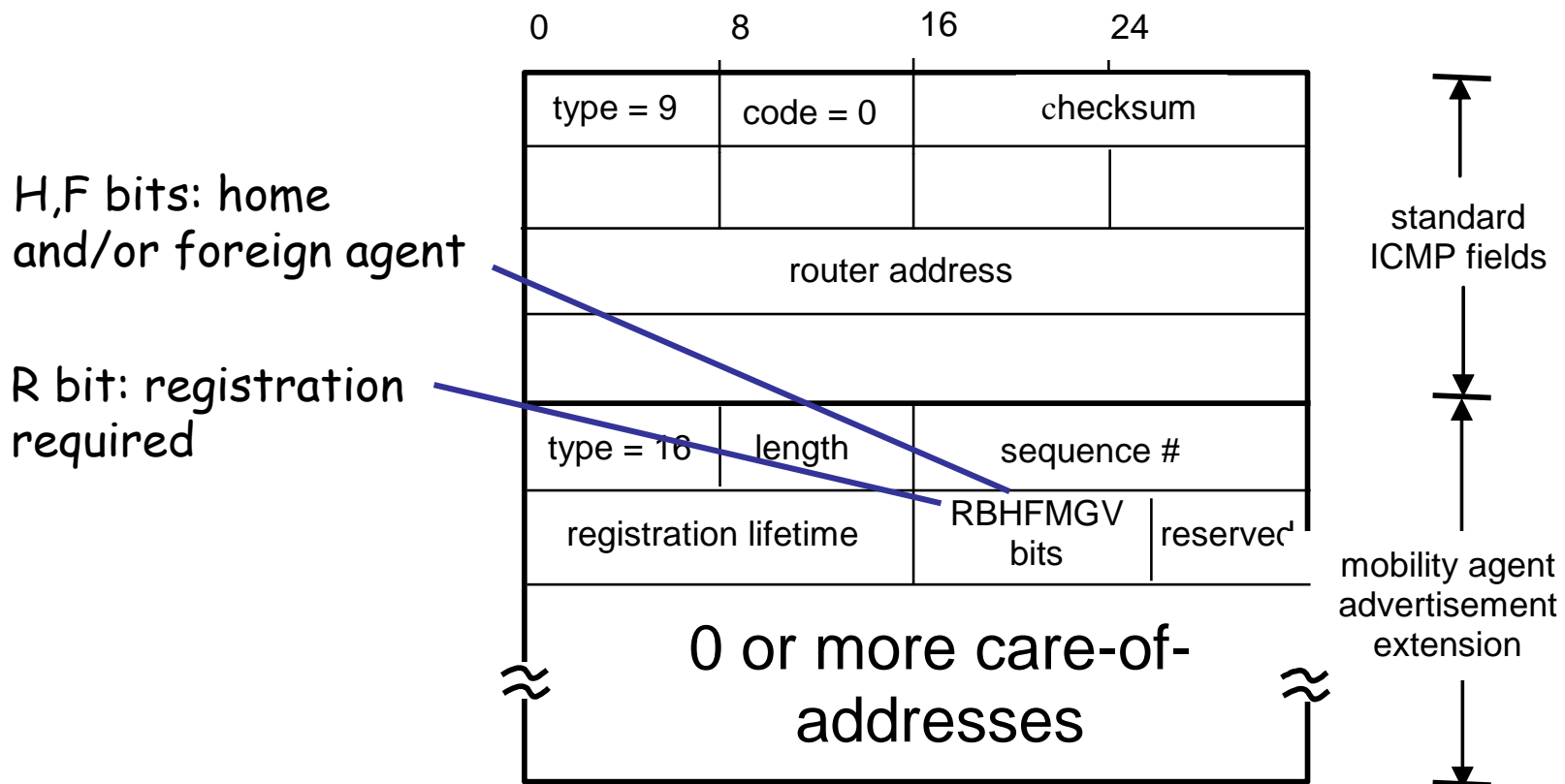


移动IP

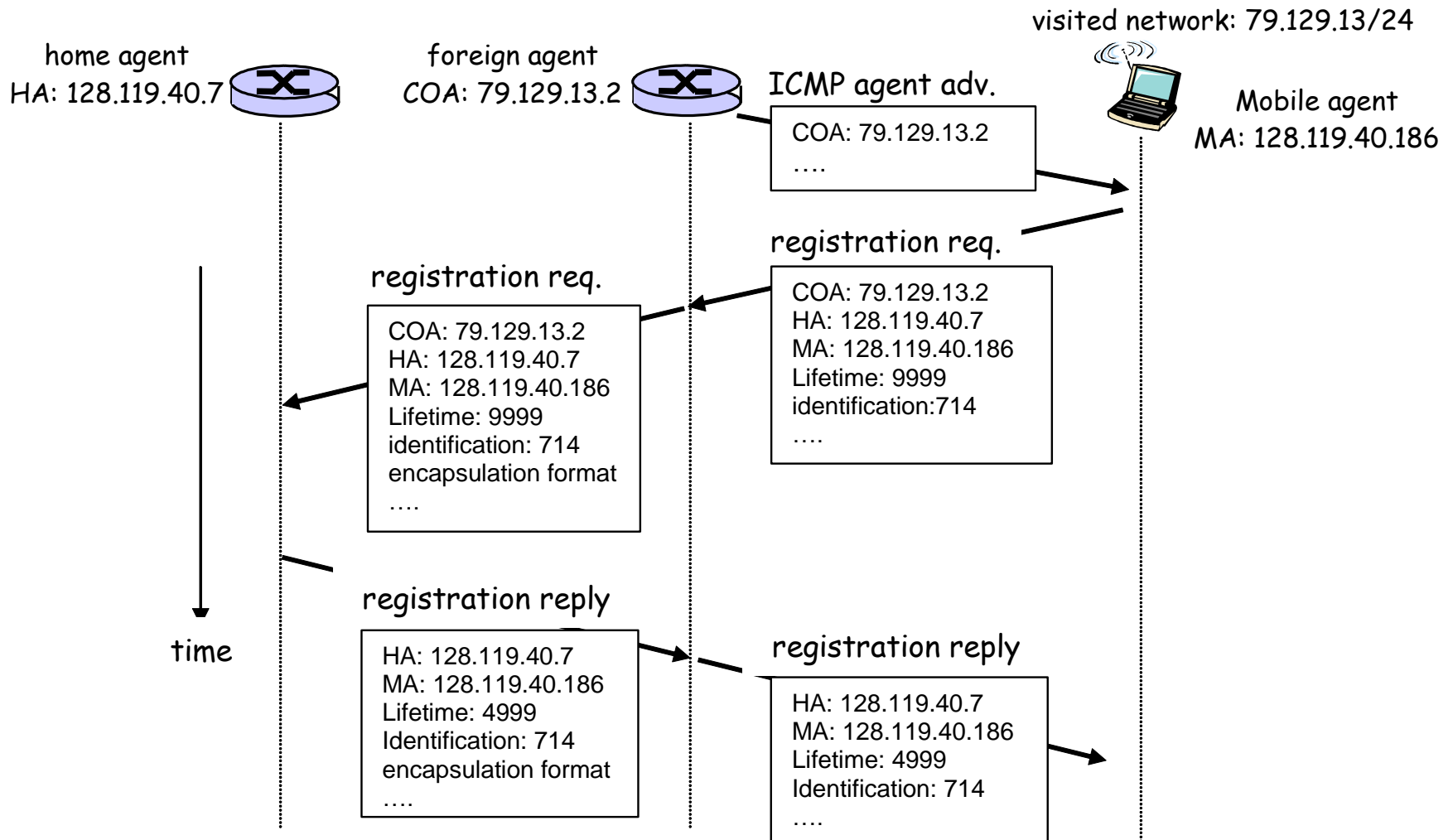
- RFC 3344, 主要部件:
 - HA, FA, FA注册, CoA (care-of-addresses), 封装
- 移动节点根据HA/FA节点的代理通告, 获得它当前的位置
 - 可以采用链路层技术或网络层技术
- 当节点移动并改变网络连接点时, 获得CoA
- 移动节点向HA注册, 建立 (CoA, MNIP) 地址绑定关系
- 建立隧道
 - 目的地址为MN的IP分组, 由HA经过隧道传送给FA或MN

移动IP: 代理发现

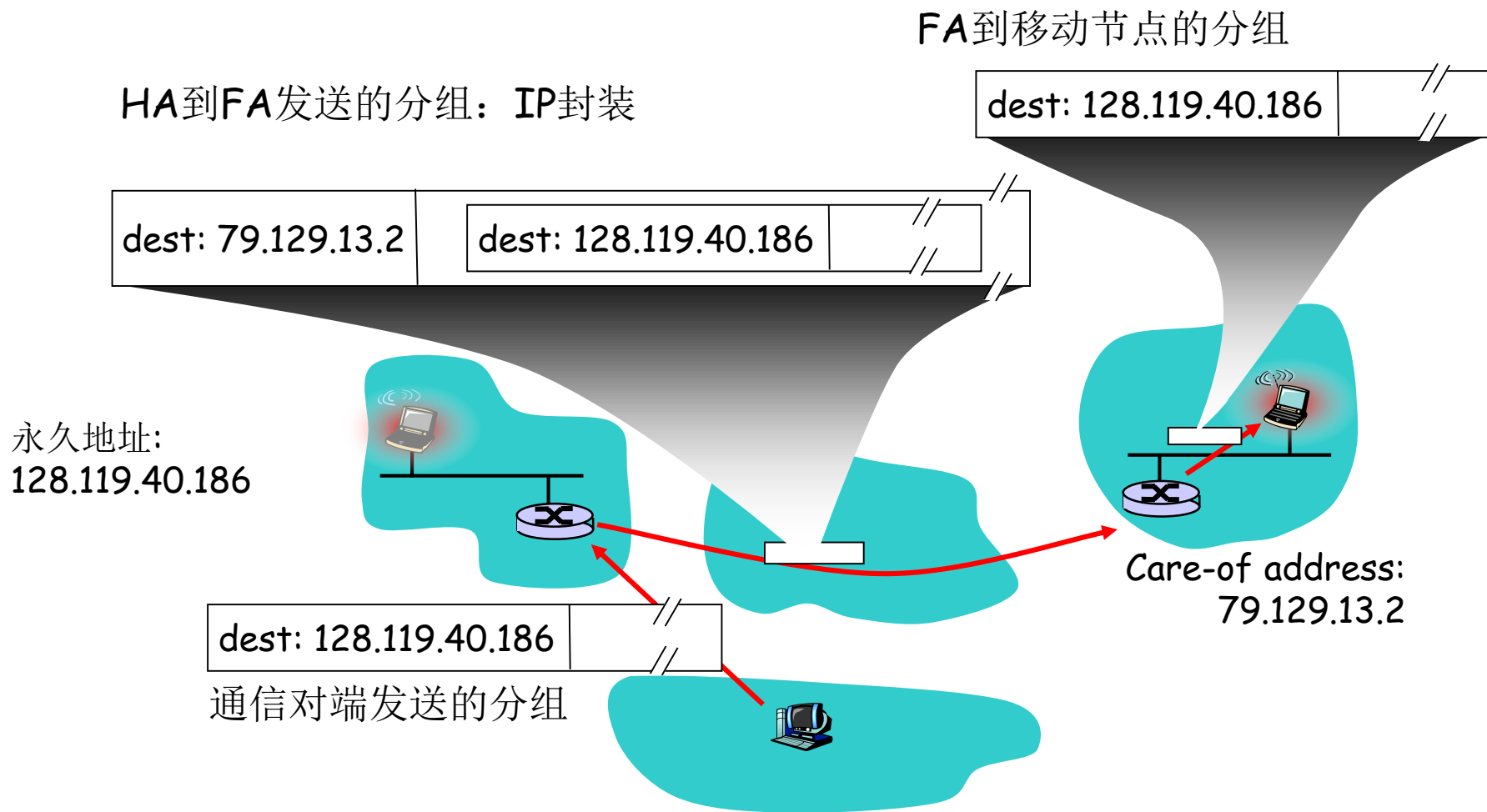
- 代理通告: FA/HA 广播ICMP 消息 (typefield = 9), 通告移动业务



移动IP: 注册过程



移动IP：间接路由



- 若HA在LAN上，问发送给本地的分组如何被HA截获？



移动IP

- 增加代理：HA、FA
- 每个移动结点需要占用2个IP
- 采用隧道技术



再议路由问题

- 路由与转发：
 - 路由器执行路由协议与算法、形成路由表及转发表
 - 最小距离路由、距离矢量路由
 - 路由器接收分组、查表选择转发端口
- IP组播及路由
- 移动主机路由
 - 网络中的路由器不移动，结点移动，采用移动IP
- 自组织网络路由
 - 网络中的所有节点都移动
 - MANET(Mobile AdHoc)
 - WSM((Wireless Sensor Networks)



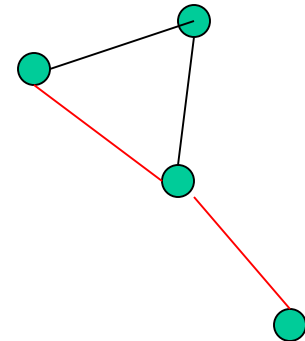
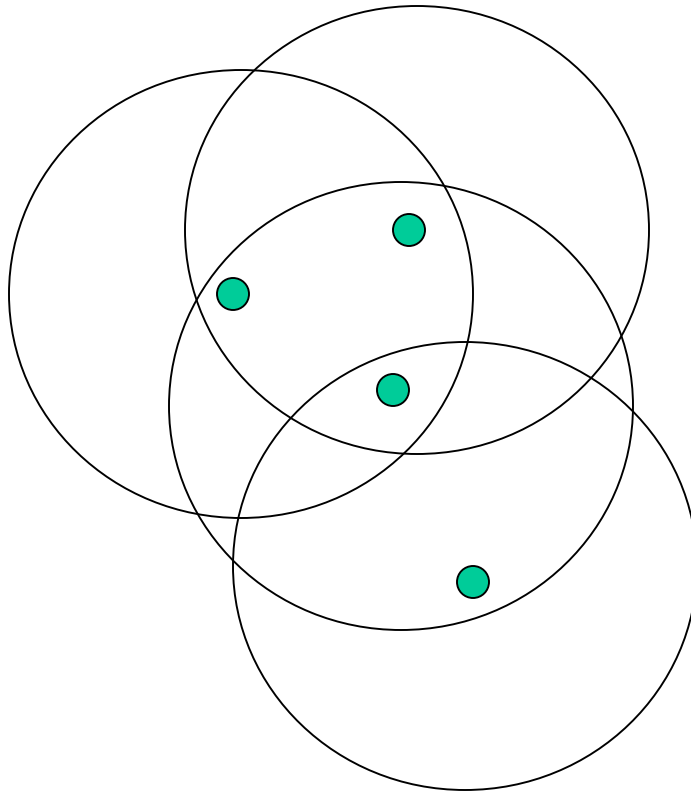
自组织（Ad Hoc）网络

■ 何谓Ad Hoc

- Ad hoc is a Latin phrase which means “for this [purpose]”.
- 由主机、移动节点组成的无线网络
- 无固定基础设施
- 网络拓扑结构可变：网络中的所有节点均可以移动
- 节点之间的路由为多跳：为便于节点移动，要求设备小型化、低功耗，因而，节点的通信能力有限，为支持更远距离的节点之间的通信，需要多跳

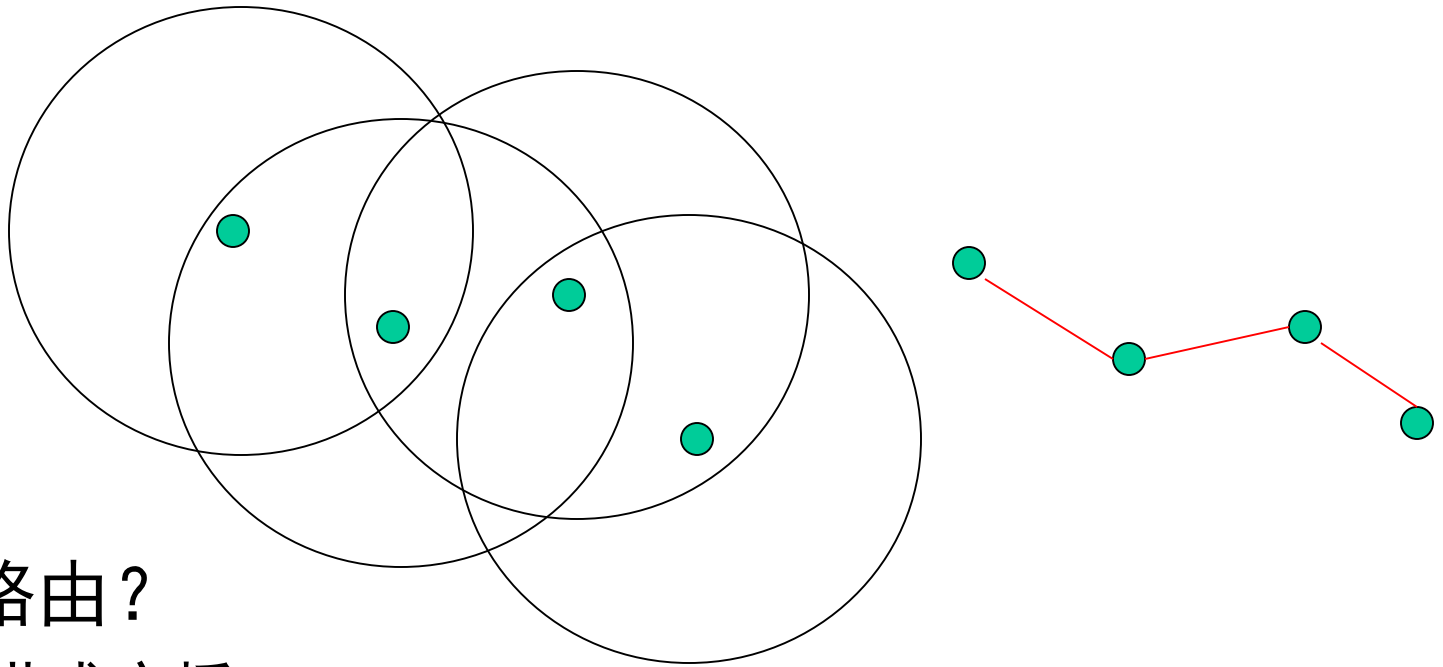
Ad Hoc Networks

- 分组传输需要经过多跳（经过多个节点的中继）才能到达目的节点



Ad Hoc Networks (MANET)

- 移动导致网络拓扑改变，要求路由变化

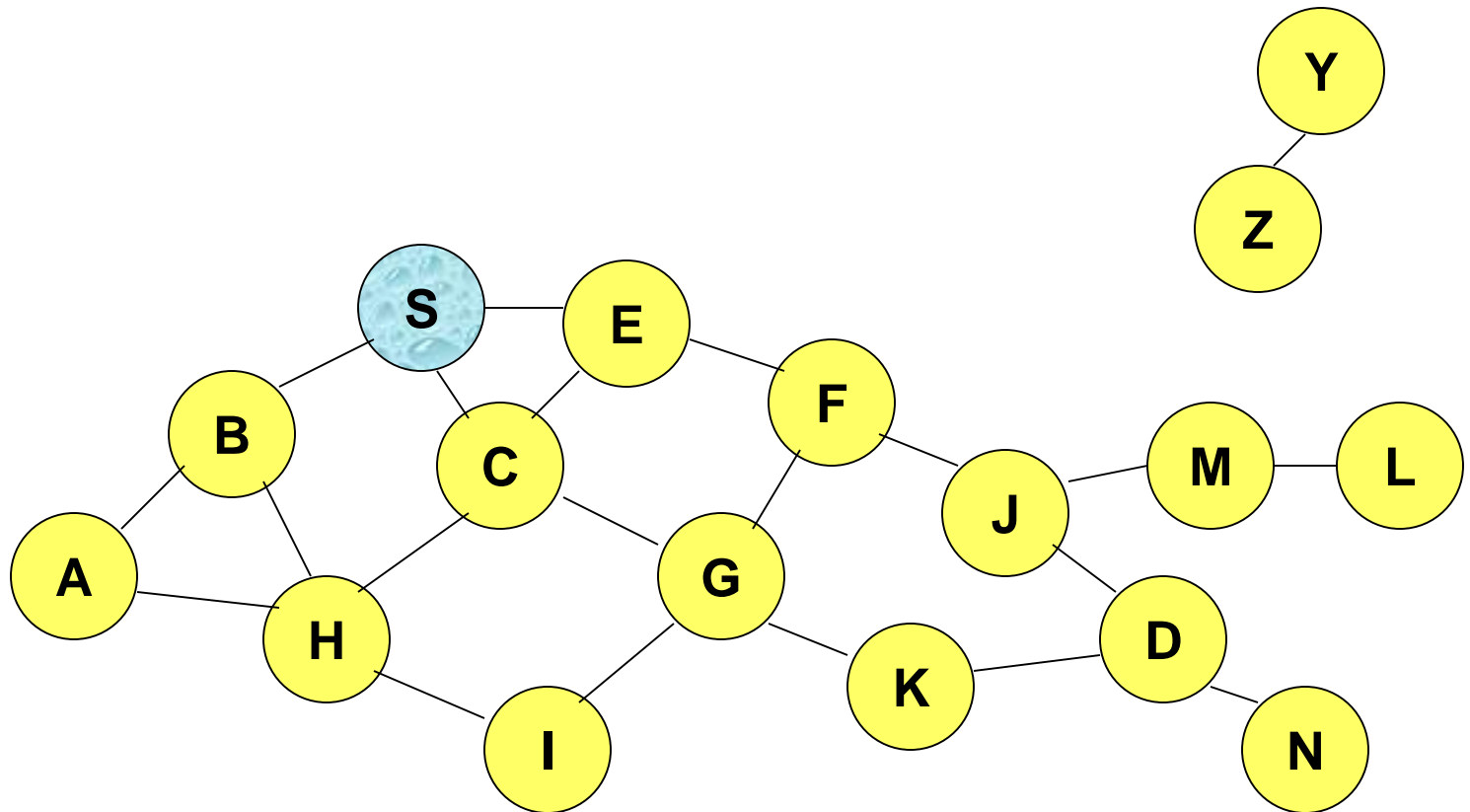


- 如何路由？
 - 泛洪或广播？
 - 最短路径或距离矢量路由？

AODV路由协议

- AODV: Ad Hoc On-Demand Distance Vector
- 当节点S要发送分组给节点D, 但没有到节点D的路径, 则启动路由发现过程 **route discovery**
- 源节点S泛洪**Route Request (RREQ)** 分组
- 其他节点收到RREQ后再广播, 同时建立它到源节点的反向路径
 - AODV假设链路是双向对称的
- 当目的节点D收到RREQ时, 发送**Route Replay (RREP)** 作为响应
- RREP沿着RREQ转发过程中所建立的相反路径转发

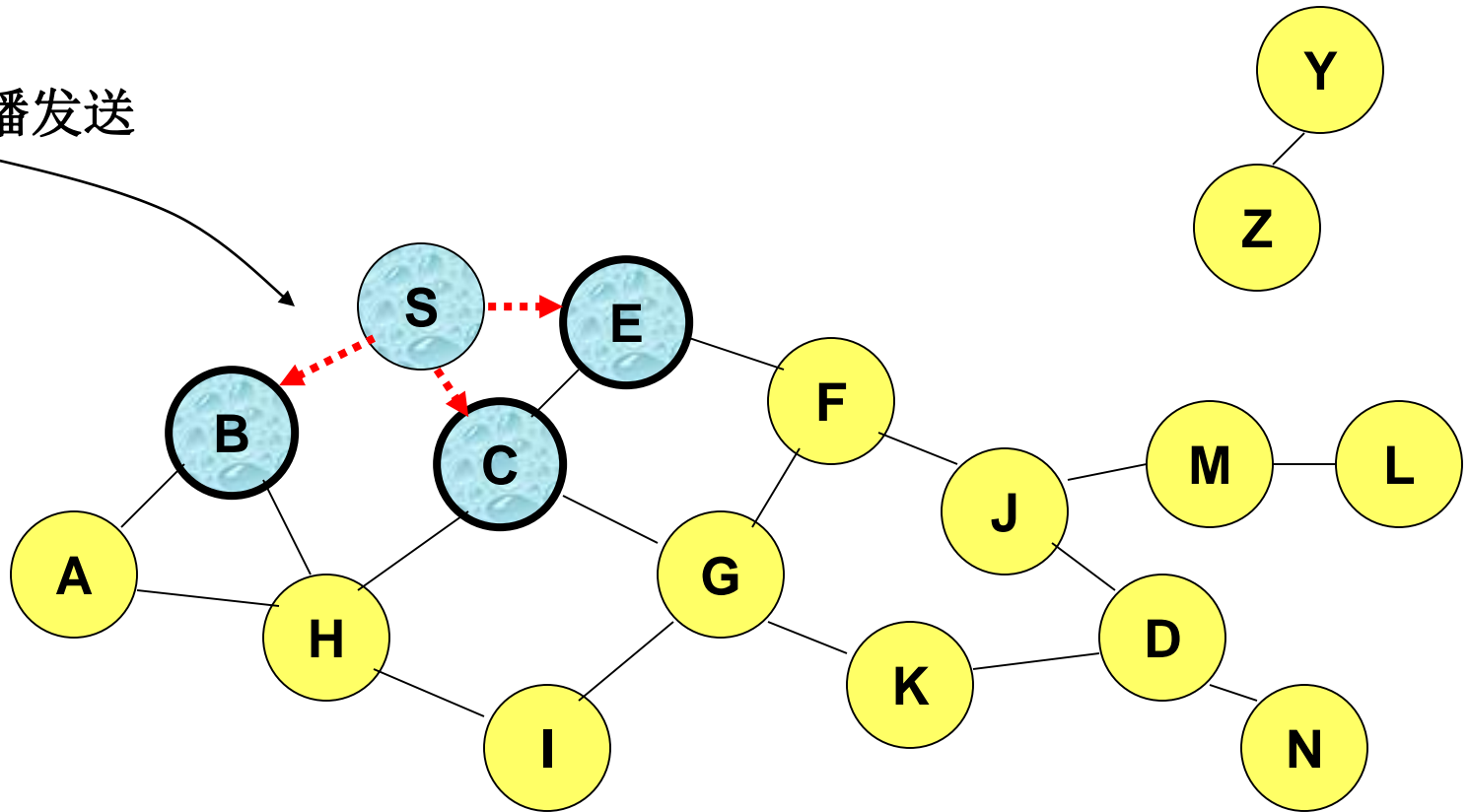
Route Requests in AODV



代表接收从**S**到**D**的**RREQ**分组的节点

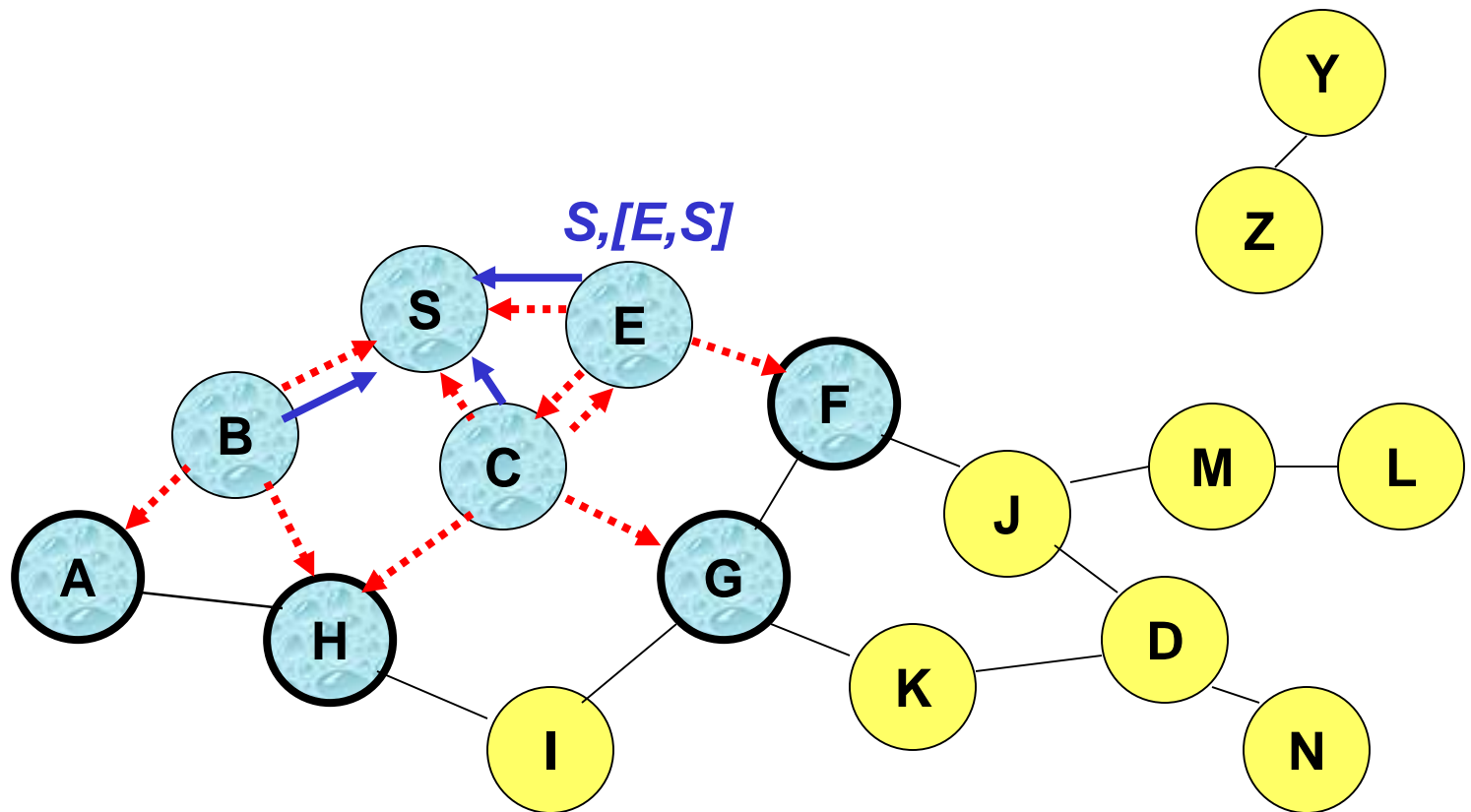
Route Requests in AODV

广播发送



.....→ 表示发送RREQ

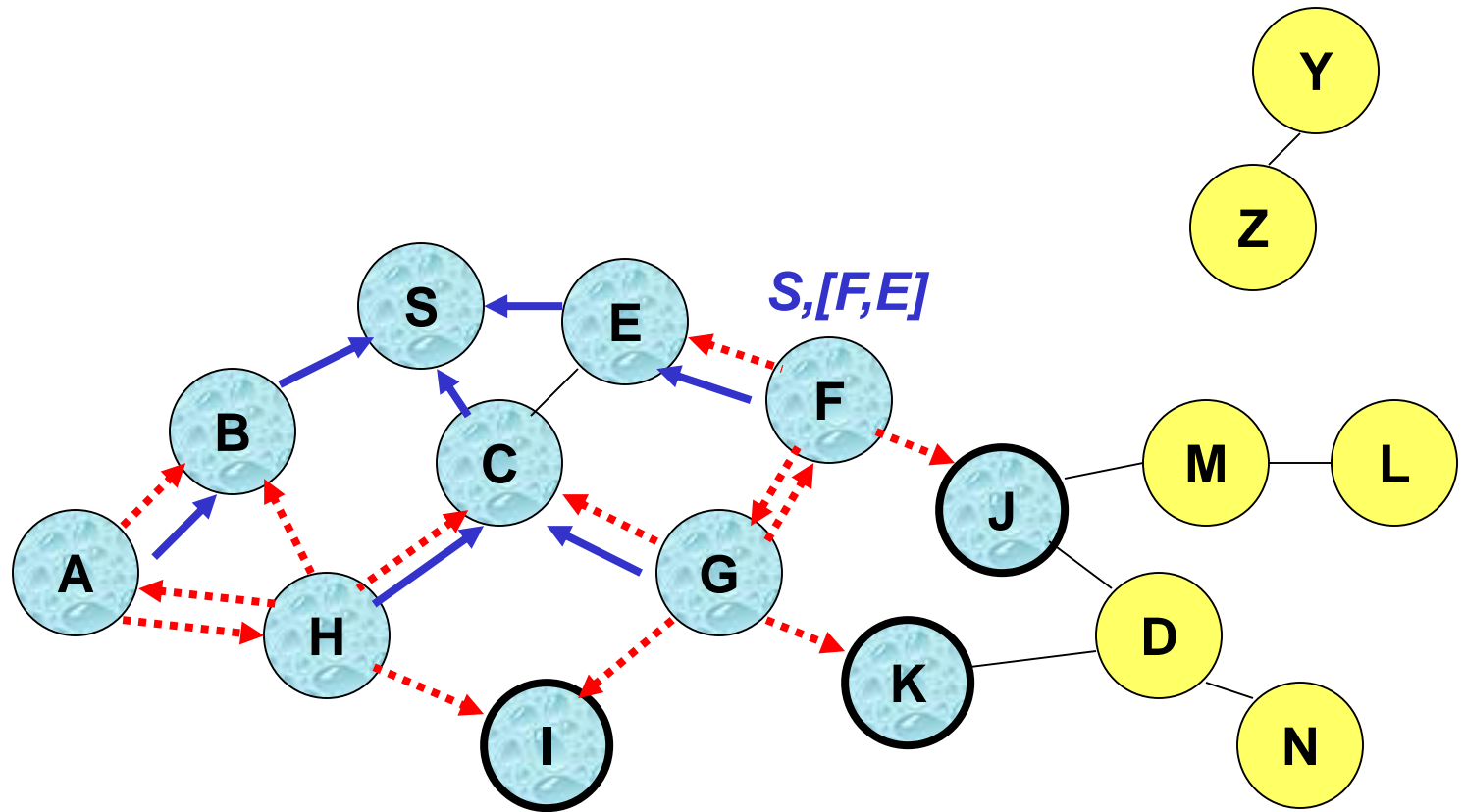
Route Requests in AODV



← 表示反向路径的链路

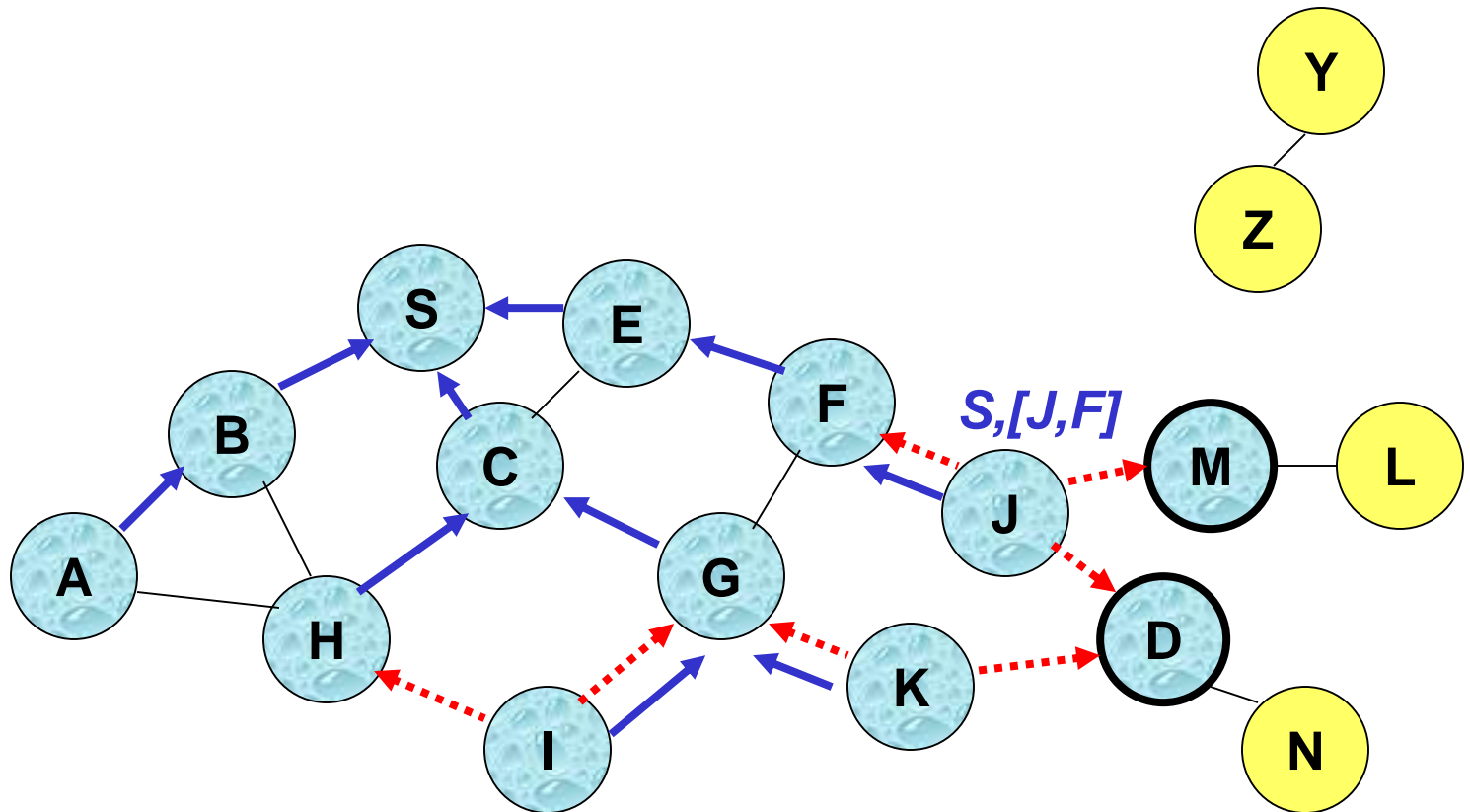
$S,[E,S]$ 表示到目的节点 **S** 经过转发链路 $[E,S]$

Reverse Path Setup in AODV

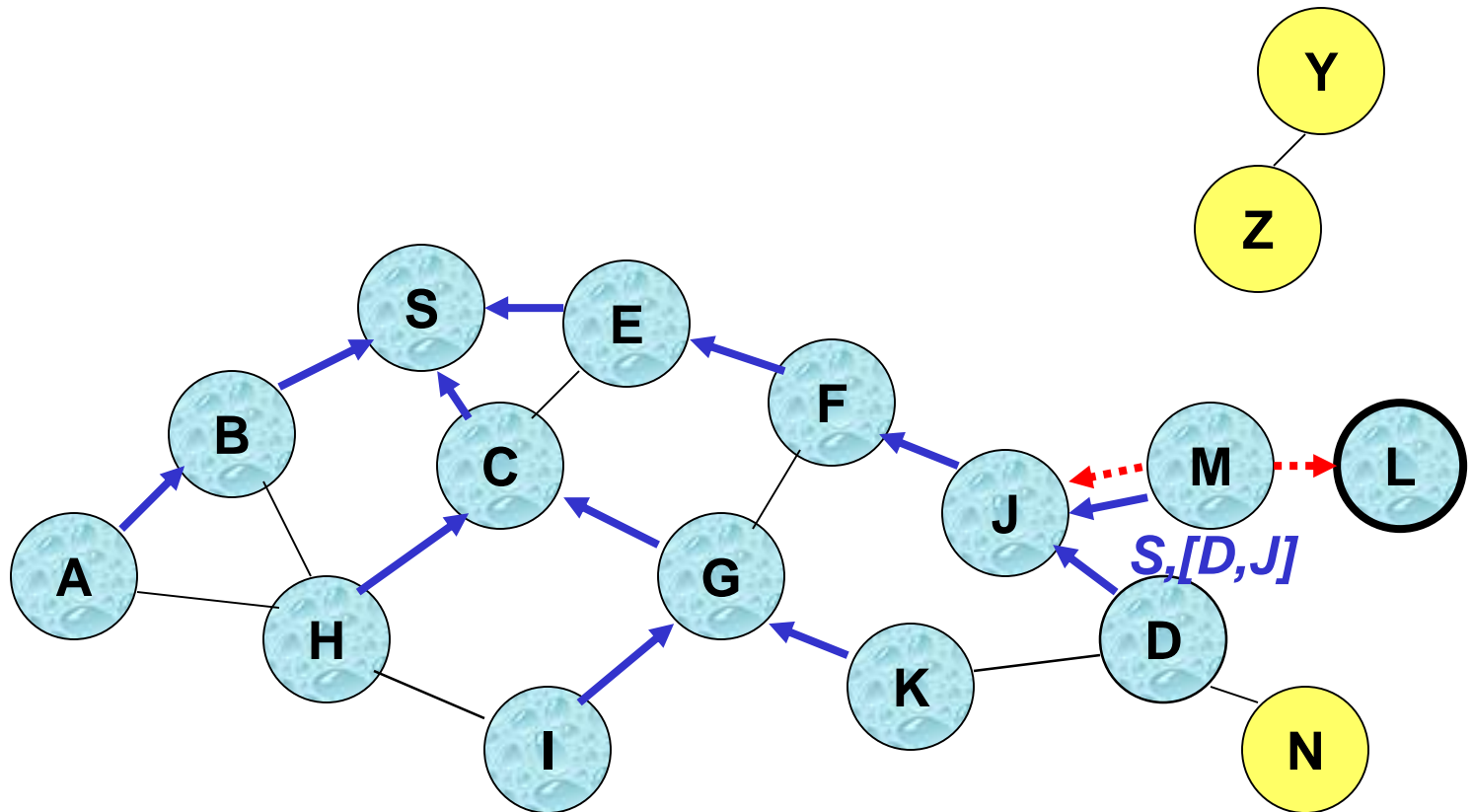


- 节点C 收到从G和H的 **RREQ**，但不再转发，因节点 C 已经转发过一次**RREQ**

Reverse Path Setup in AODV

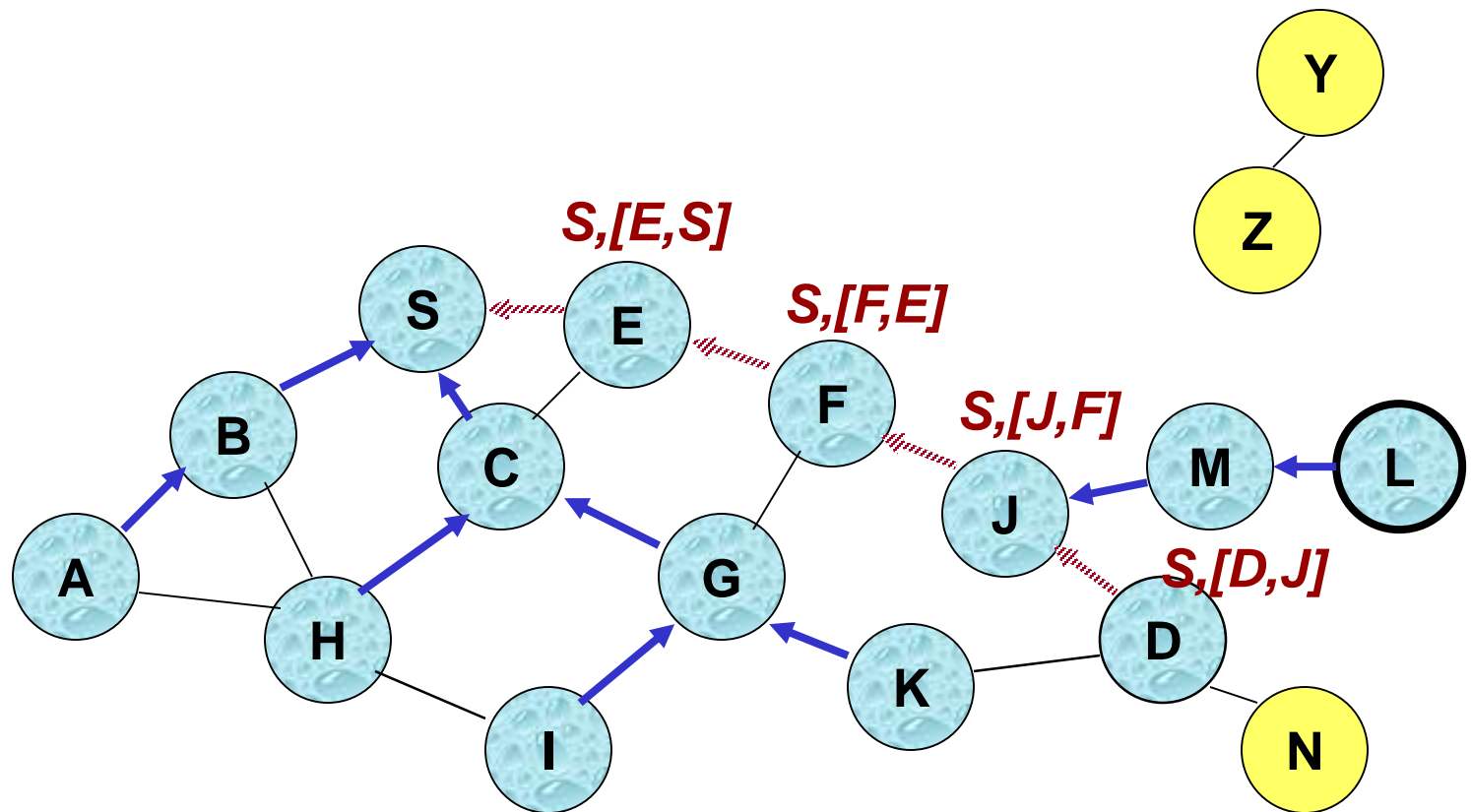


Reverse Path Setup in AODV



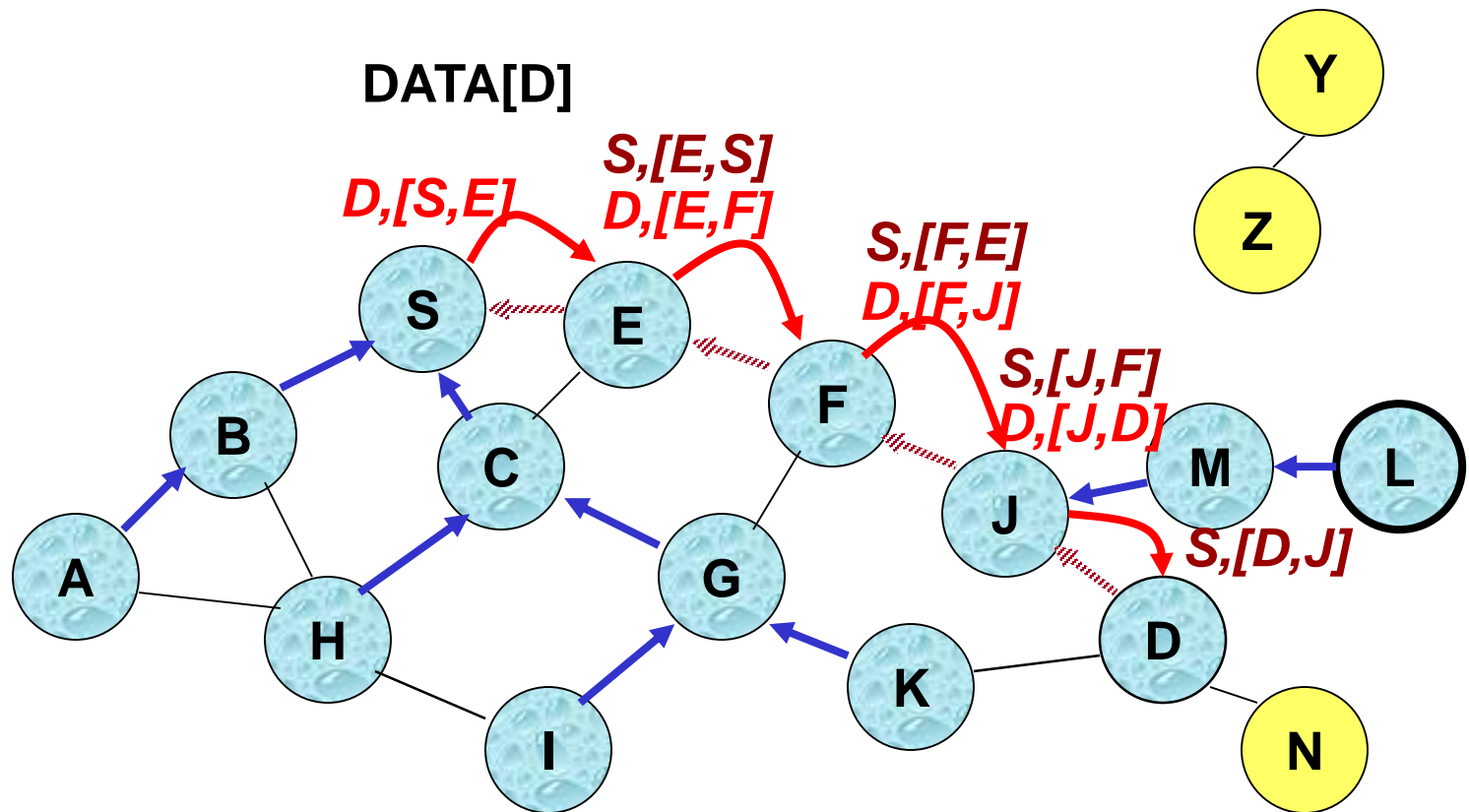
- 节点D不再转发RREQ, 因 D为RREQ 的目的节点

Route Reply in AODV



 表示RREP选择的链路

Forward Path Setup in AODV



当RREP沿反向路径传输时，则建立了D到S的前向链路



代表前向路径上的链路

在DATA中不含路由信息，各节点根据路由表转发分组

Route Reply in AODV

- 为确定路径信息新旧并选择更新的，定义目的地序列号DSN *destination sequence numbers*，在RREQ、RREP、RERR中含有DSN
- 若节点S到节点D发送一个新的RREQ，就分配一个更大的DSN
- 中间节点有到目的节点D的路由，若其DSN
 - 比RREQ中的小，则不能发送RREP
 - 比RREQ中的大，则可以发送RREP

Timeouts

- 在广播RREQ过程中所建立的路由表项，维持了D到S的一条反向路径是临时的，若路由表项超时，则被删除
 - 所设置的超时时间应足够长，以保证可返回RREP
- 路由表项所维持的S到D的一条前向路径，如果在`active_route_timeout`内该路径未被使用，则删除
 - 在路由表的某端口上若`active_route_timeout`时段内没有发送数据，即使该路径是有效的，也要删除该路径表项
 - 降低保存路由表的内存开销

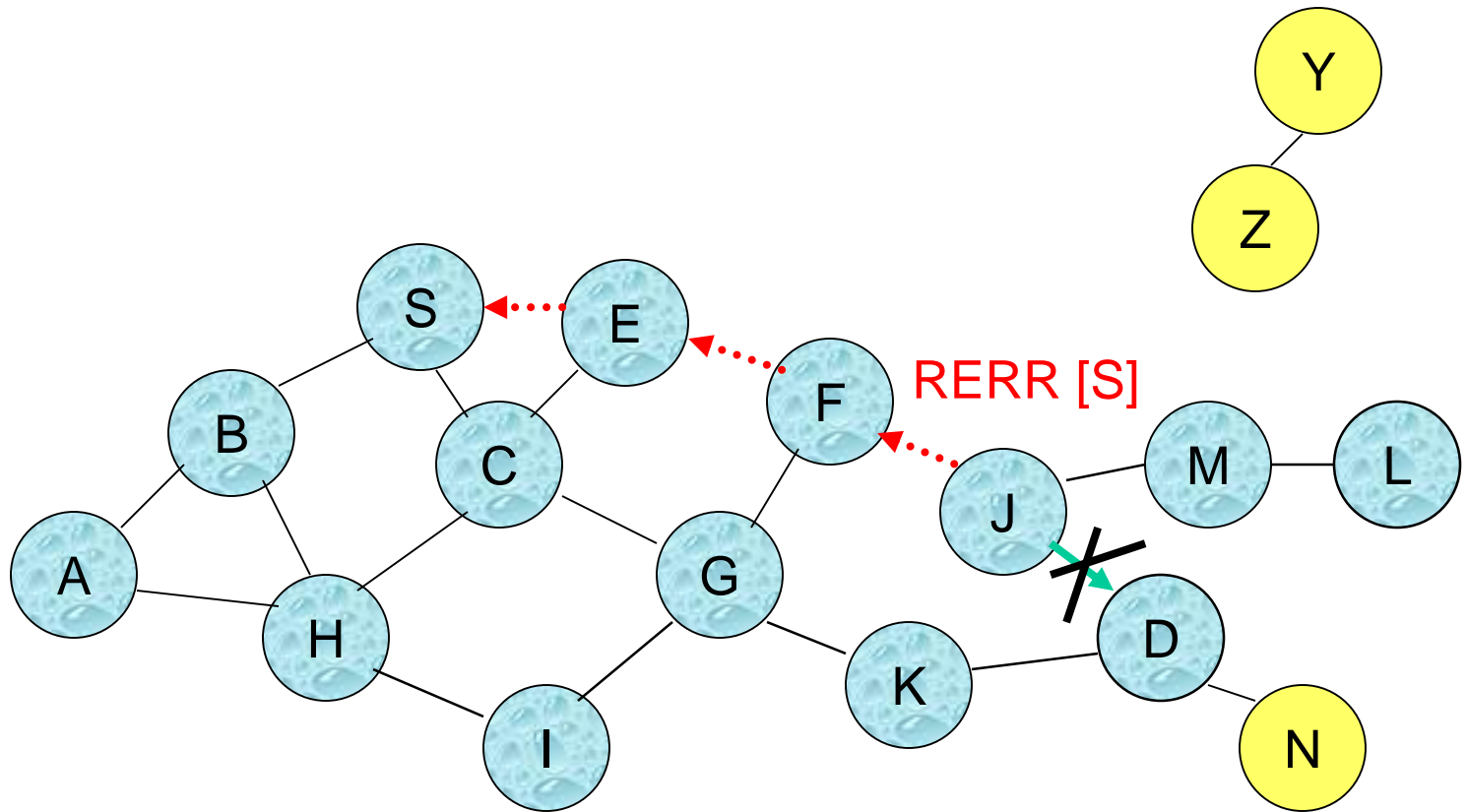
Link Failure Reporting

- 如果在 *active_route_timeout* 间隔，通过路由表的相应端口转发过分组，则认为对应的节点X是一个相邻节点并处于激活状态
- 当节点路由表的下一跳节点链路中断，则通知所有激活节点
- 传播链路失效消息，采用Route Error (RERR)，同样，要更新RERR中的DSN

Route Error

- 当节点X不能在链路 (X, Y) 上转发从节点S到节点D的分组时，则产生一个RERR消息
- 节点X根据缓存的目的节点D的DSN，其DSN加1
- 在RERR中，含有增加的DSN
- 节点S接收到RERR，启动一次新的到节点D的路由发现过程，其DSN再加1
- 节点D收到RREQ，若DSN更大，则设置新的DSN

Route Error (RERR)



- 当J在J-D链路上转发分组失败时，则DSN+1并向S发送路由错误分组RERR；收到RERR的节点，转发RERR并删除路由表项
- S接收到RERR, 发送RREQ（其中的DSN加1）启动到节点D的新的路由发现过程

AODV: 小结

- 在分组头中不含路由信息，降低了头部开销
- 节点维护一张路由表，每一路由表项设置一个定时器，表示路由是否过期
- 一般情况下，每个节点为每个目的节点维持下一跳地址
- 即使网络拓扑不变，不用的路由也会因过期而被删除



本章小结

- 网络层的功能：为传输层提供与路由器的数量、类型和拓扑结构无关的服务；统一编址；网络拥塞控制、**保证服务器质量**、网络互联
- 路由器：网络层互联设备，实施路由与转发
- 网络互联
 - 地址分配：IP地址，三种编址方式；DHCP、NAT与专用地址
 - 分组传送：ARP，分段与重组
 - 隧道技术
- 路由：距离矢量路由算法与RIP；最小距离路由算法与OSPF
- 网络控制：超时、差错恢复、状态报告、拥塞控制
- IP组播及路由：IGMP、组播路由
- 移动主机路由：代理（HA、FA）通告、注册、隧道
- 自组织网络路由