Chapter 4

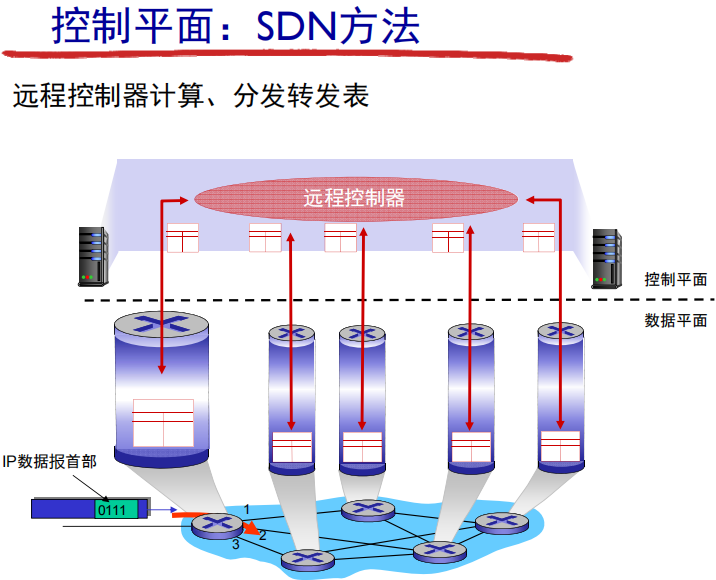
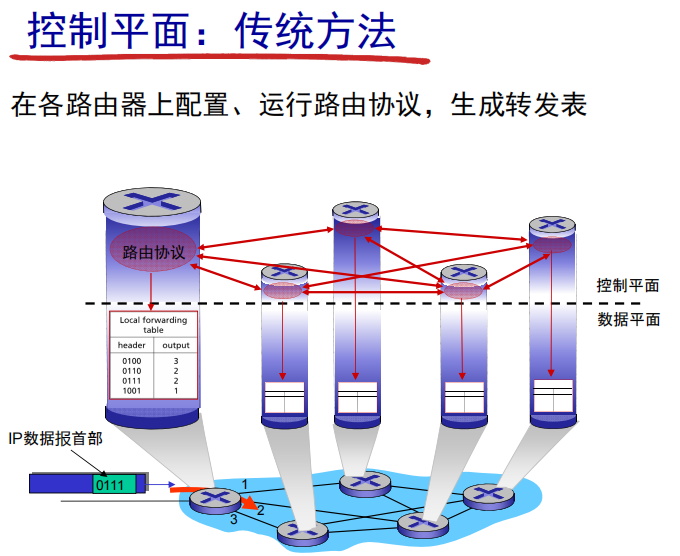
1. 网络层功能:转发和路由

* 转发：当一个分组到达某个路由器的输入链路时，该路由器必须将其移动到合适的输出链路。（单个路由器，数据平面）
* 路由：决定分组从源主机，到目的主机的传输路径。（整个网络，控制平面）
* 路由的两种方法：

传统的路由协议：分布式，所有路由器中实现

软件定义网络software defined network(SDN): 集中式，远程服务器中实现

2.



3. 分组交换机：

链路层交换机：根据链路层字段值作转发决定的分组交换机。路由器：根据网络层字段值作转发决定的分组交换机。

4.网络层提供的服务：尽力而为服务,能否交付，是否按序，时延均无保证

虚电路网络：提供连接服务。如帧中继和ATM

* 根据虚电路号（VC号）转发分组，该路径上每段链路的号码，每条链路上的VC号可能不同
* 可同时建多条虚电路，服务可靠：到达顺序与发送顺序一致
* 路由器维护连接状态信息，只要该路由器创建新的VC，其转发表中就增加一项；终止一个VC，其转发表中就删除对应项

虚电路的三个阶段：

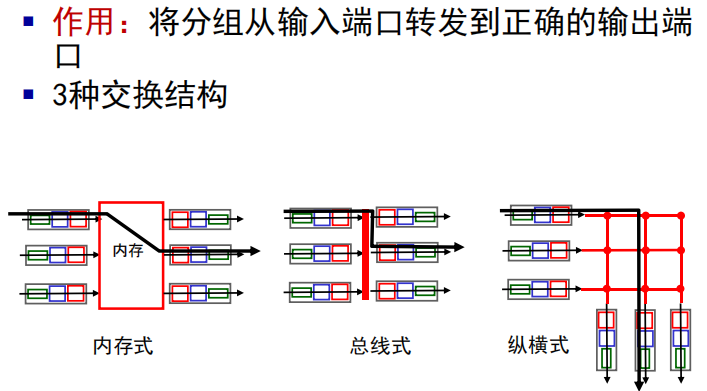
* 虚电路建立：在发送方与接收方之间建立一条虚电路，即决定所有分组要通过的一系列链路与路由器，并为每条链路确定一个VC号
* 数据传送：沿该虚电路传输数据分组
* 虚电路拆除：由其中一方通知其网络层终止该虚电路；通知网络另一侧的端系统呼叫结束，并更新路径上每台路由器中的转发表。

数据报网络：提供无连接服务。

* 分组使用目的主机地址转发，相同源和目的地的分组可能采用不同的路径
* 传输过程：发送方给要发送的分组加上目的端系统地址，并送入网络；经过若干中间路由器转发分组，直到目的地。
* 路由器转发方法：根据到达分组的目的地址在转发表中查询，找到相应的输出链路接口，并将分组转发出去
* 路由器查表方法：用目的地址前缀与转发表的前缀比较匹配

5.路由器结构

交换结构



排队：

输入端口：输入端口速率之和 > 转发速率

输出端口: 转发速率>输出端口速率

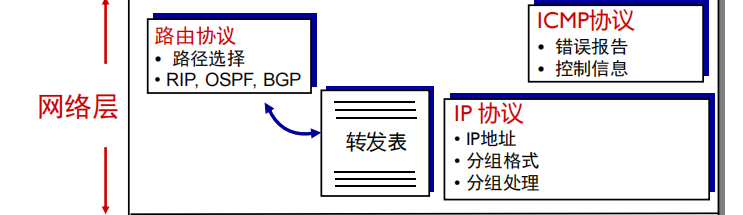
分组调度：在输出端口排队的分组中选出一个发送

先来先服务FCFS：简单。加权公平排队WFQ：在具有排队分组的不同端到端连接之间公平地共享输出链路。

6. 线头HOL (head-of-the-line)阻塞： 输入队列中后面的分组被位于线头的一个分组阻塞(即使输出端口空闲的)，等待交换结构发送。



7.网络层结构

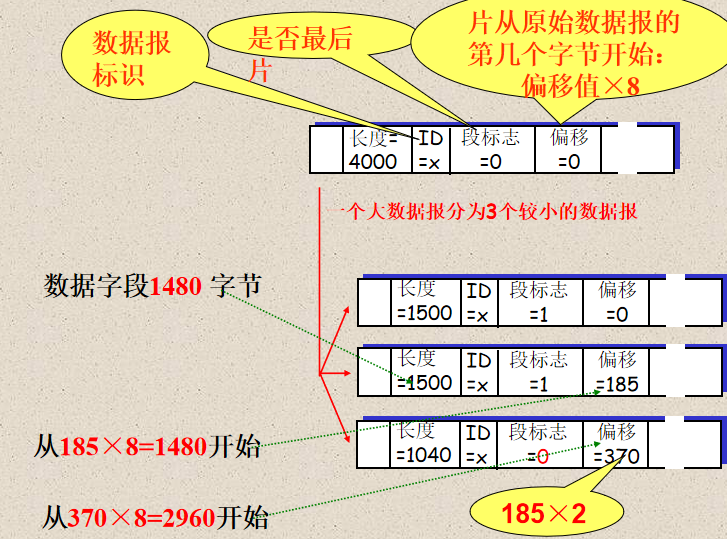


8.IP：网际协议

该协议不可靠、尽力而为、无连接

9. IP分片和重新组装

* 网络链路帧有最大传输单元（MTU）长度，不同链路长度不同。
* 分段：大IP 数据报被分割成几个传输
* 重装：在目的主机组装。

e.g. 

10.IP地址：互连网中主机分配的一个唯一地址。IPv4长度为32比特，IPv6扩展到128位

11.DHCP动态主机配置协议

* 基于UDP的应用层协议
* 作用: 主机自动获取IP地址，地址续约，地址重利用，但移动用户无法保持连接
* DHCP步骤:

主机广播发送“DHCP discover” 报文[可选]

DHCP 服务器广播响应“DHCP offer”报文[可选]

主机选择服务器，广播发送“DHCP request”报文

DHCP服务器广播返回“DHCP ack”报文

12. ISP如何获得IP地址块

ICANN: 因特网名字和编号分配机构

• 分配IP地址

• 分配域名

• 管理DNS根服务器

13.网络地址转换NAT

作用:

▪ 只使用一个公有IP地址，节约地址

▪ 内网地址变动，不用通知外网

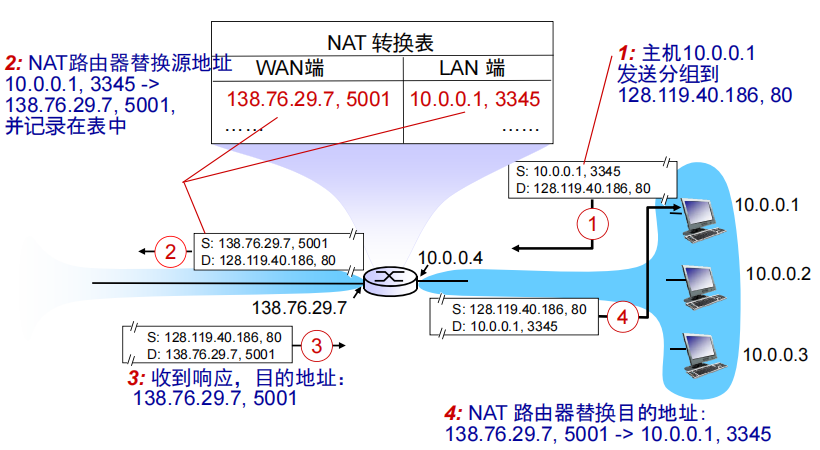
▪ 更换ISP，无需更改内网

▪ 内网结构不可见，更安全

实现: NAT路由器上:

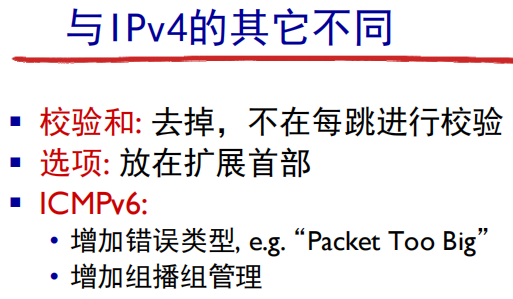
▪ 出去的分组: 替换 (源IP地址, 端口 #) 为(NAT IP地址, 新的端口#)，并记录，远方主机将(NAT IP地址, 新的端口#) 作为目的地址，返回应答

▪ 进来的分组: 逆向替换

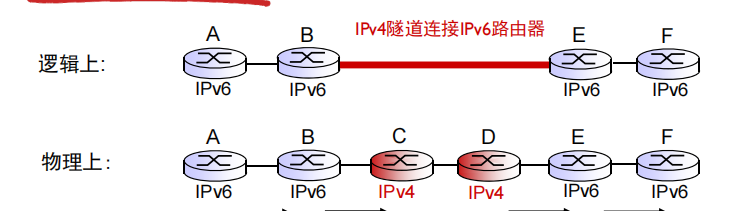


14.IPV6

服务类型: 低时延、高吞吐量、高可靠性



路由器无法同时升级,同时存在IPv4和IPv6的路由器，如何通信? 将IPv6分组封装到IPv4分组，在IPv4路由器间传输



15.默认路由器

与主机直接相连的一台路由器，又叫第一跳路由器，主机发送一个分组时，都先传送给它的默认路由器

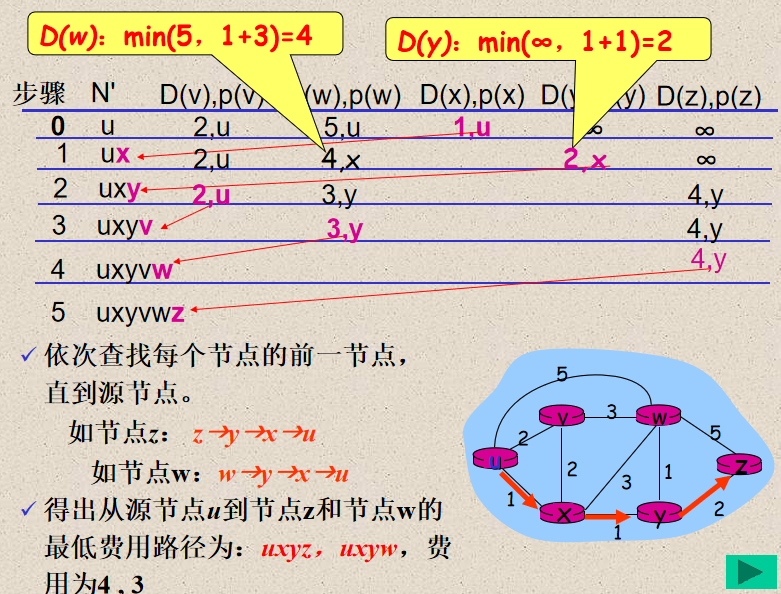
源路由器：源主机的默认路由器目的路由器：目的主机默认路由器

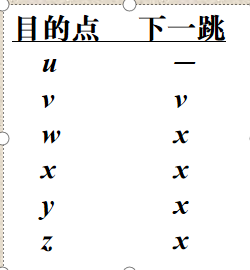
从源主机到目的主机的选路：从源路由器到目的路由器的选路。16. 常用选路算法

链路状态算法LS

距离向量算法DV

LS：

* 前提条件：已知网络拓扑和所有链路的费用，作为算法的输入。（全局链路状态已知）
* 获取方法：每个节点向网络中广播链路状态分组（含有它所连接的链路的费用）。由链路状态广播算法实现。最终使所有节点都有一个相同且完整的网络视图
* Dijkstra最低费用路径算法：计算从某节点（源节点，如u）到网络中所有其他节点的最低费用路径。一种迭代算法：经第k次迭代后，可知道到k个目的节点的最低费用路径。
* 节点u的转发表：



* 默认路由：即将“下一跳”相同的项合并为一项，目的节点用“\*”表示。优先级最低，转发分组时，当找不到对应表项时，才使用默认路由。
* 算法复杂性O()
* 缺陷：时间复杂度高，计算量大，路由器需要维护整个链路状态，开销大，并可能产生“振荡”
* 避免振荡：避免所有的路由器同时运行LS算法：比较合理。即使路由器以相同周期运行LS算法，但在每个节点上算法的执行时刻不同

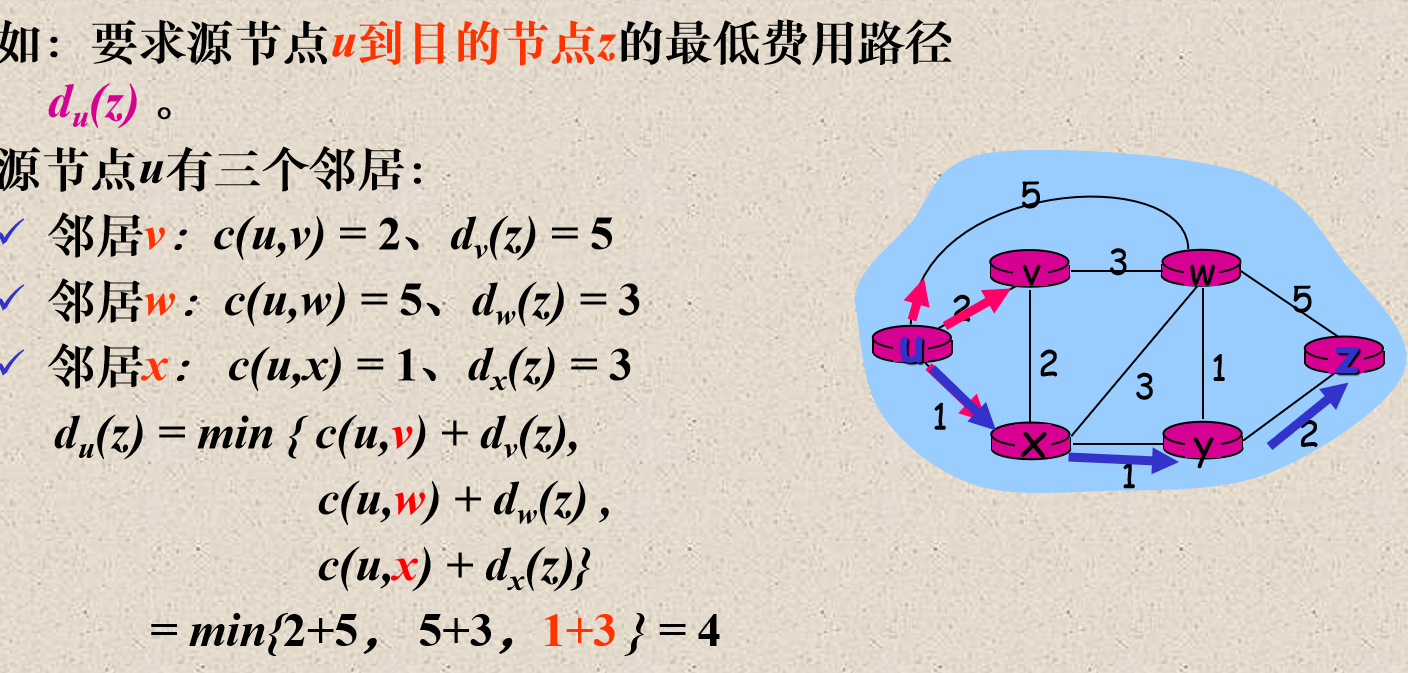
DV（距离向量算法）：

* 迭代的、异步的和分布式的算法

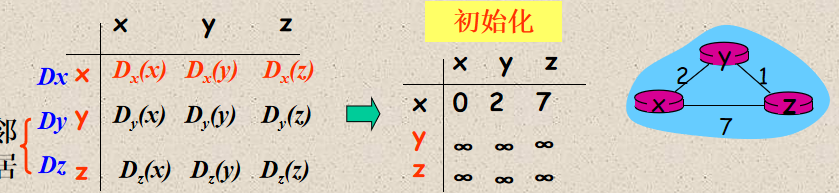
分布式：每个节点都从其直接相连邻居接收信息，进行计算，再将计算结果分发给邻居。迭代：计算过程一直持续到邻居之间无更多信息交换为止。自我终结：算法能自行停止。异步：不要求所有节点相互之间步伐一致地操作

* Bellman-Ford方程表示 dx(y) = minv{c(x,v)+ dv(y)}

其中，dx(y)：节点x到节点y的最低费用路径的费用。

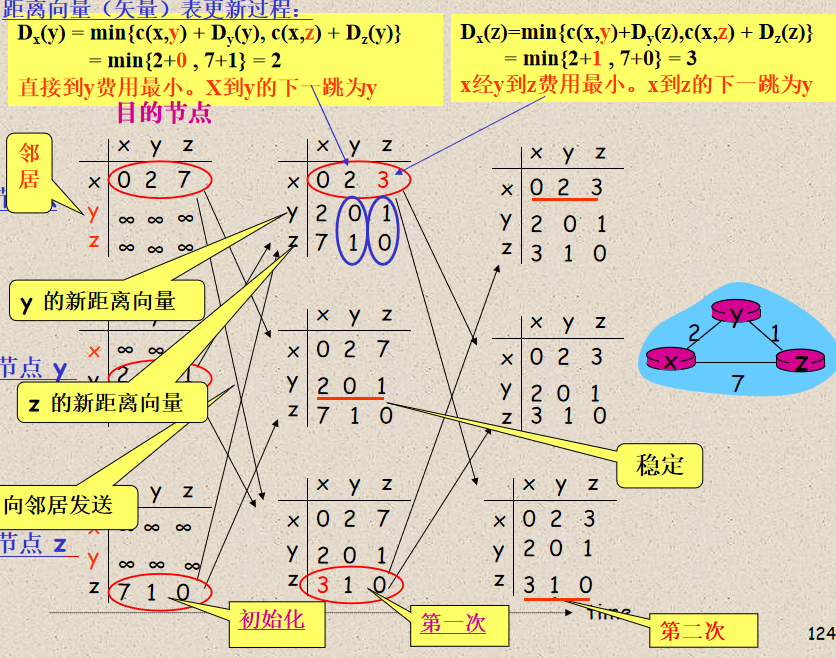


* 节点x选路表

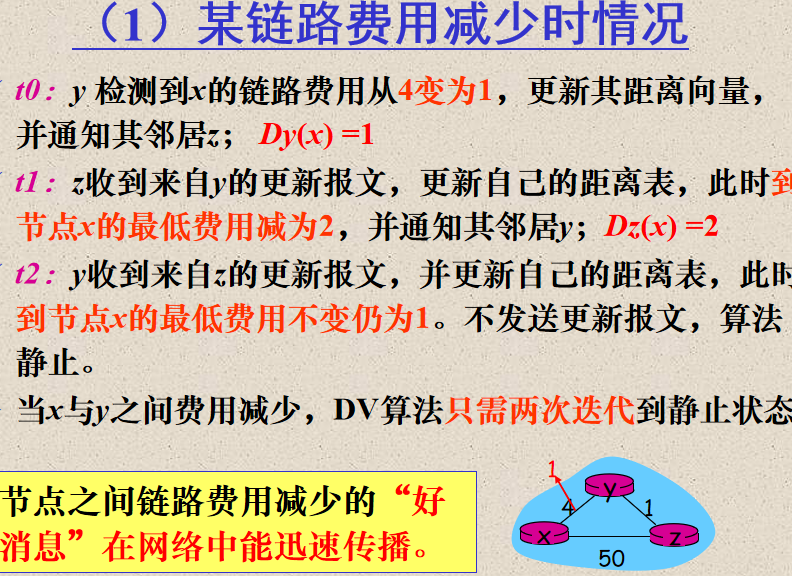


* 每个节点不断向邻居发送其距离向量拷贝；当节点x收到一个邻居v的新距离向量，先保存，并用下列公式更新自己的距离向量： Dx(y) = minv{c(x,v)+ Dv(y)} 若距离向量发生改变，将新的距离向量通知给邻居。何时更新：直接相连的链路费用变化，或收到邻居的新距离向量，当距离向量不再变化，算法静止，此时Dx(y)收敛到dx(y)，即得到节点x到节点y的最低费用路径，
* 算法描述：对每个节点x（1）初始化：（2）更新自己的距离向量 （3）重复执行（2），直到没有更新的距离向量发出。

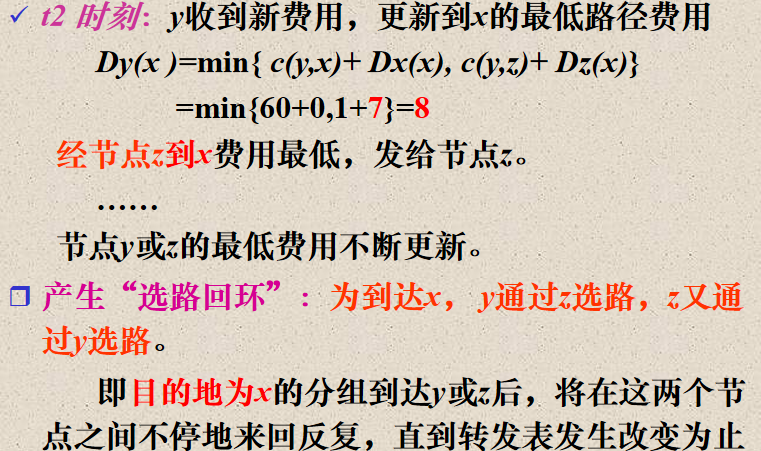
发现直接相连的链路费用变化，或收到邻居的新距离向量，更新自己的距离向量，在经每个邻居v到目的点y的距离向量中，选择一个最小值作为当前新距离向量，并将该邻居v作为节点x转发表中到目的节点y的下一跳路由器v\*(y)。若距离向量发生改变，向其邻居发送新距离向量



* 节点之间链路费用减少的“好消息”在网络中传得快。



* 节点之间链路费用增加的“坏消息”在网络中传得慢



* 毒性逆转：如果 z 通过 y 选路到达目的地 x，则 z 告诉 y，它到 x 的距离是无穷大，局限是三个或更多节点的回路，毒性逆转技术检测不到， 因此， “不可计数问题”也未解决。

17. LS算法与DV算法比较

DV算法：每个节点只与邻居互相交流，得到邻居的新费用，并告知邻居自己的当前最低费用LS算法：每个节点与所有其他节点广播交流，只告知与其直接相连链路的费用

* 报文复杂性：LS算法：知道网络每条链路的费用，需发送O(nE)个报文；当一条链路的费用变化时，必须通知所有节点DV算法：迭代时，在两个直接邻居之间交换报文，收敛时间受许多因素影响；当链路费用改变时，只有该链路相连的节点的最低费用路径发生改变时，才传播已改变的链路费用。
* 收敛速度：LS算法：需要O(nE)个报文和O(n2)的搜寻。DV算法：收敛较慢。可能会遇到选路回环，或计数到无穷的问题。
* 健壮性：考虑一台路由器发生故障、操作错误或受到破坏LS算法：路由器向其连接的一条链路广播不正确费用。路由计算基本独立（仅计算自己的转发表），有一定健壮性。DV算法： 一个节点的计算值会传递给它的邻居，并间接地传递给邻居的邻居。一个不正确的计算值会扩散到整个网络。

18.其他路由算法

* 热土豆选路算法：路由器尽可能快地转发输出分组。向任意不拥塞的输出链路转发分组，不管目的地在哪里。、
* 网络流算法：把分组流量看成是一个网络中源和目的地之间的流。将选路问题形式化为数学上的受限优化问题。
* 电路交换选路算法：每条链路资源(如缓冲区、部分链路带宽)都要保留给每个要经过该链路的连接

19.层次路由

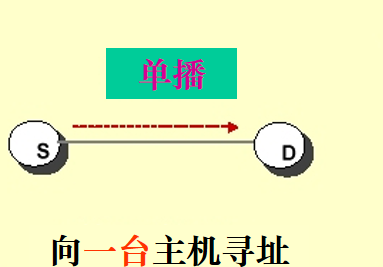
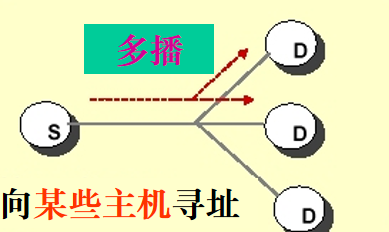
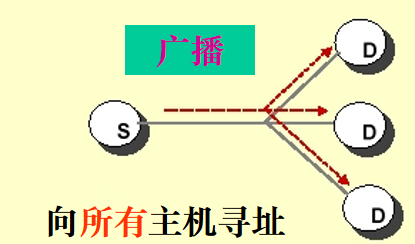
按区域或自治系统的形式组织路由器。将一个大的系统划分成若干小系统（自治系统），自治系统之间再互连。同一个AS内的路由器运行相同的选路算法

* 自治系统AS：按区域划分的系统。每个AS由一组在相同管理者控制下的路由器组成
* 自治系统内部选路协议 ：在一个自治系统内部使用的路由协议

RIP选路信息协议：基于距离向量的路由协议OSPF开放最短路径优先：采用Dijkstra最短费用路径算法，是一种链路状态协议。特点：安全，允许相同代价路径 (RIP只能一条)，支持多播，支持层次结构

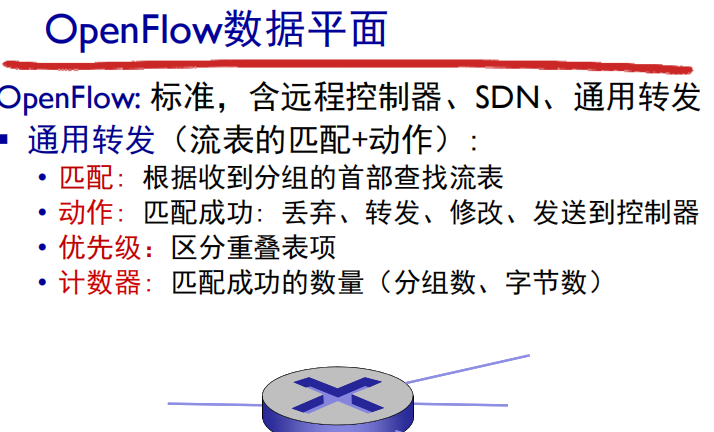
* 自治系统间选路协议：在多个自治系统之间使用的路由协议。BGP边界网关协议：基于DV的路由协议。
* 网关路由器：互连各AS，负责转发目的地在本AS之外的分组，将本AS内的分组转发到另一个AS的路由器

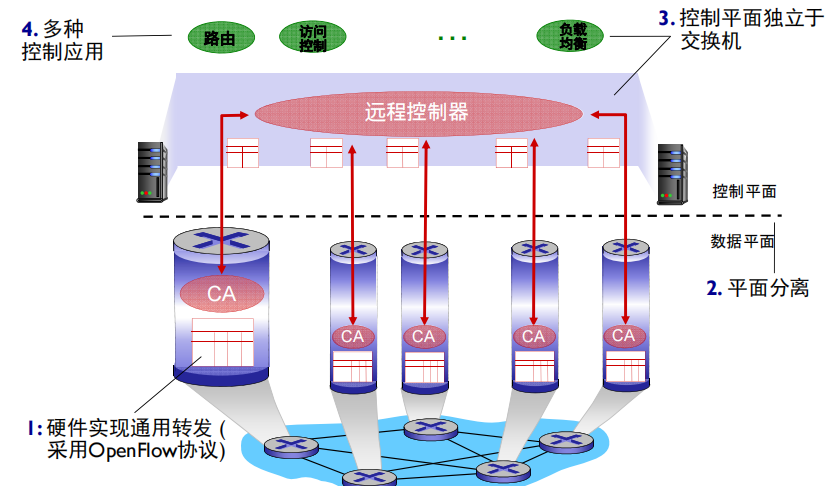
20. 几种通信方式

21.SDN

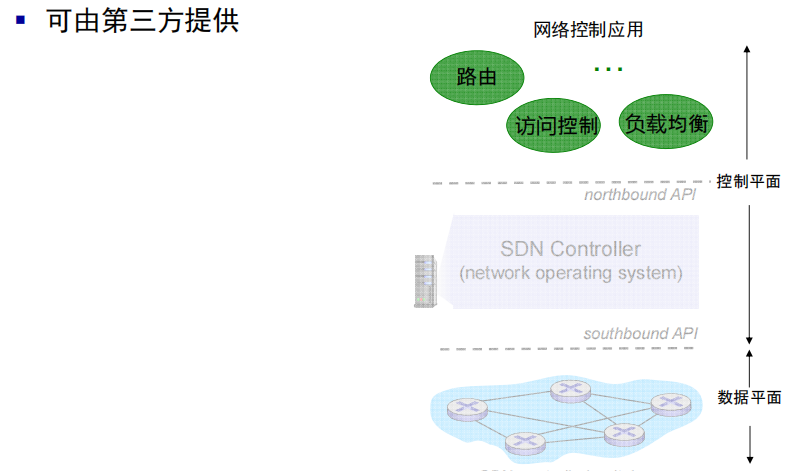
* Open flow->通用转发，运行于SDN控制器和交换机之间，基于TCP

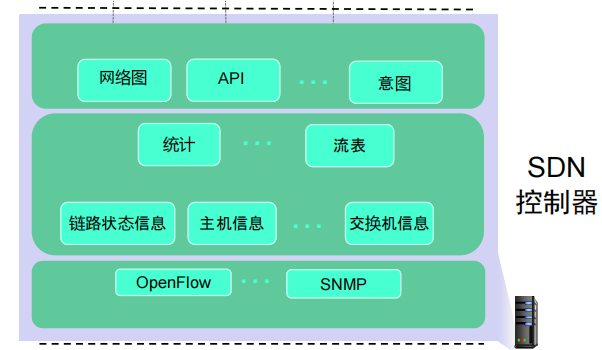












22.域间路由 vs. 内部路由

