

计算机网络层

功能

- 路由routing
 - 规划出一条从发送方到接收方的路由（路径）
 - 要求是最短的
- 转发forwarding
 - 路由器从一个端口接受，选择另一个端口发送

routing:每一个路由器形成一个转发表 forwarding table吗，里面每一个IP对应一个端口

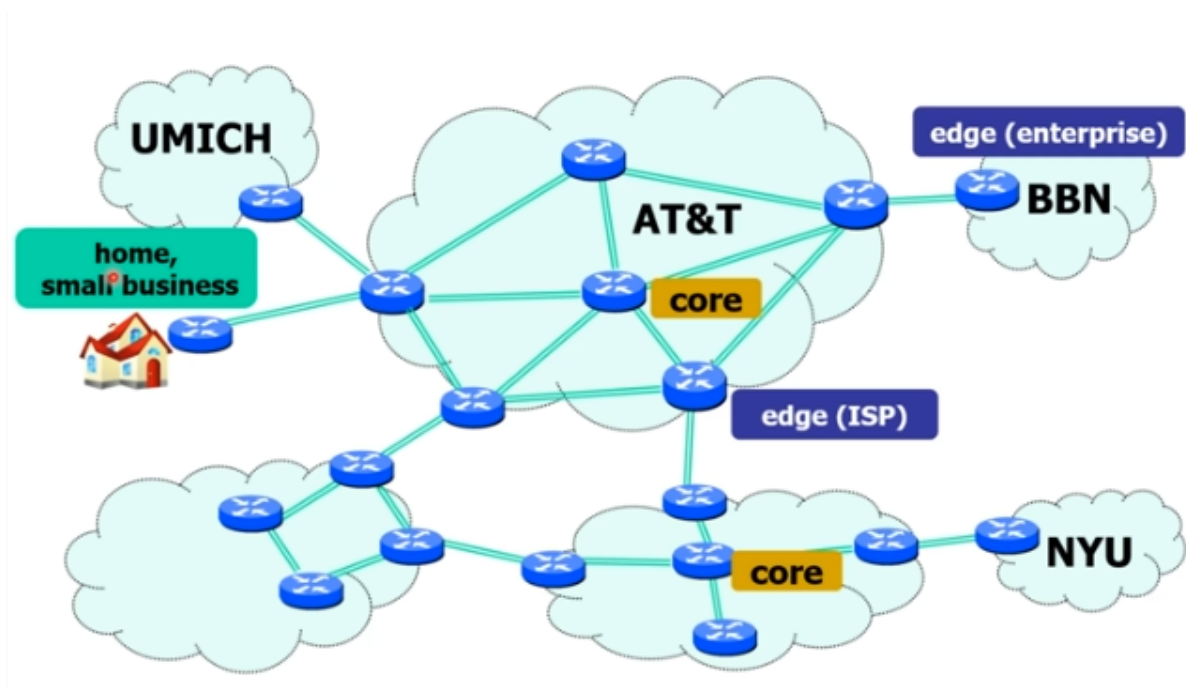
forwarding:存储并转发，路由器里面有一个缓冲区，会有排队的现象

网络服务模型

- 单独数据包
 - 保证送到
 - 保证送到的延迟不高于多少
- 数据流
 - 是否一定是按序送到
 - 最小的带宽保证
 - 数据包之间的间隔

IP路由器

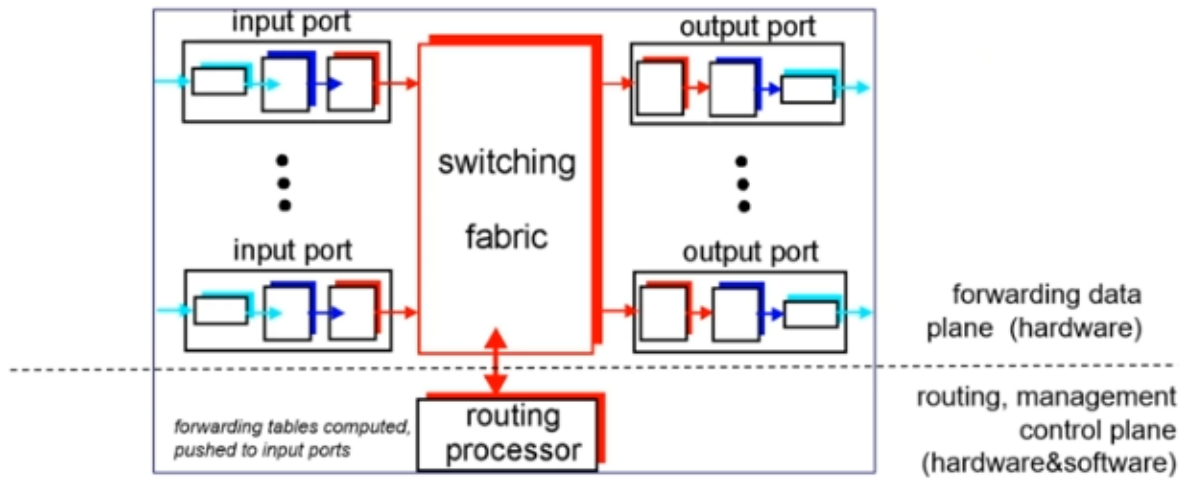
路由器的容量=端口数*端口的速率



三种路由器：core，edge，small business，区别是容量不同

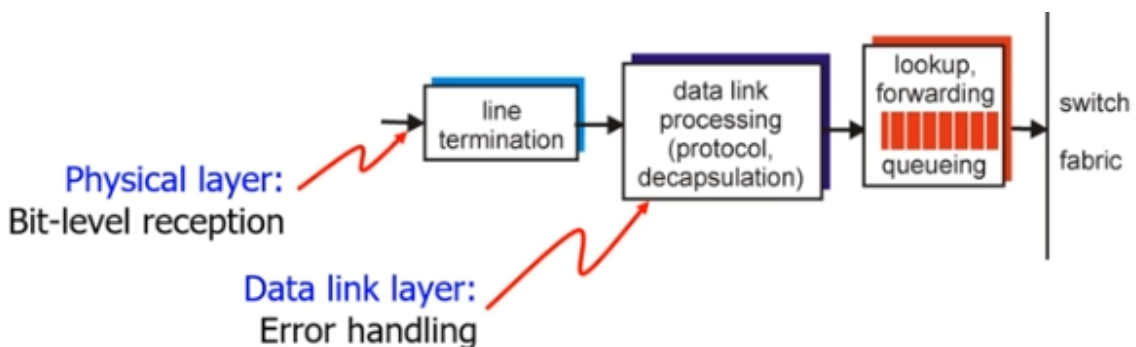
Two key **switch** functions:

- Run **routing** algorithms/protocol
- **Forwarding** packets from incoming to outgoing link



路由器分为两个部分，下面由routing，上面forwarding
routing processor由硬件和软件计算线路，做出转发表
switching fabric完全是硬件，按照转发表进行存储转发

输入端口



Tasks

- Receive incoming packets (physical layer stuff)
- Update the IP header
 - TTL, Checksum, Options and Fragment (maybe)
- Lookup the output port for the destination IP address
- **Queuing:** if packets arrive faster than forwarding rate into switch fabric

物理层—bit—bit接收

链路层把数据切成帧，加上帧头，然后有错误处理等等

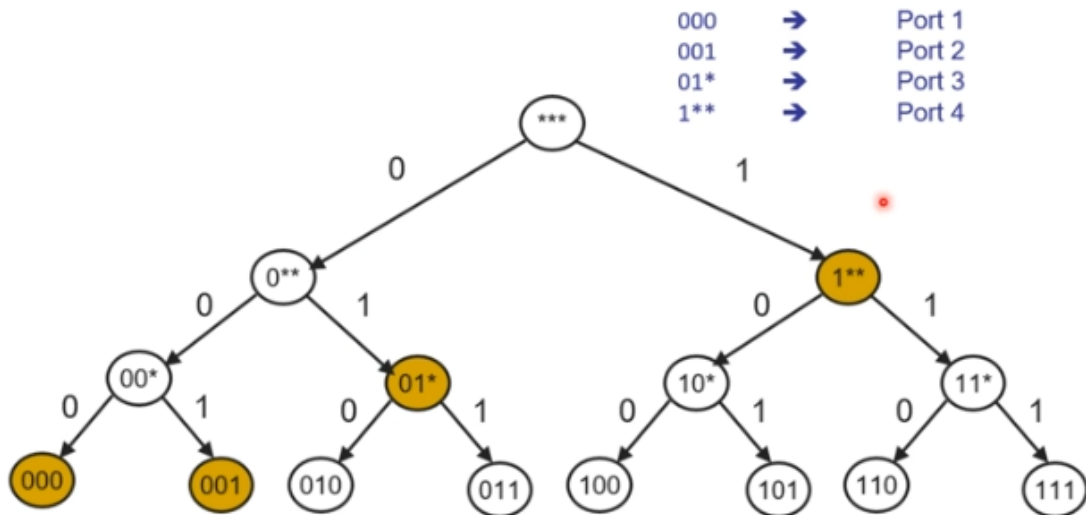
IP层：

- 接收报文并且解封装查看IP报文头
- 更新IP报文头，报文头有一些字段每一跳是需要更新的
- 查看转发表进行转发
- 如果接受的速度超过发送的速度，排队

转发表

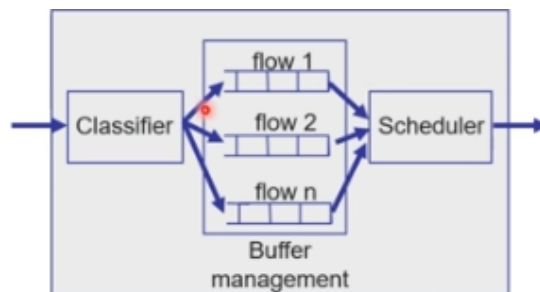
要求处理的速度非常高，那么就要求一个好的routing算法

- 不能把每一个IP都分别放在一个表项里面，那样转发表太大了，我们把一组有相同前缀的IP地址对应一个端口，放在表里面（最长前缀匹配）
- 用树状的结构，避免 $O(n)$ 查找



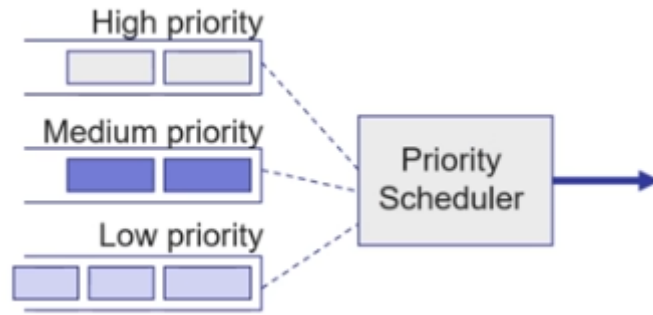
输出端口

- 最简单的：先进先出
- 现在用的



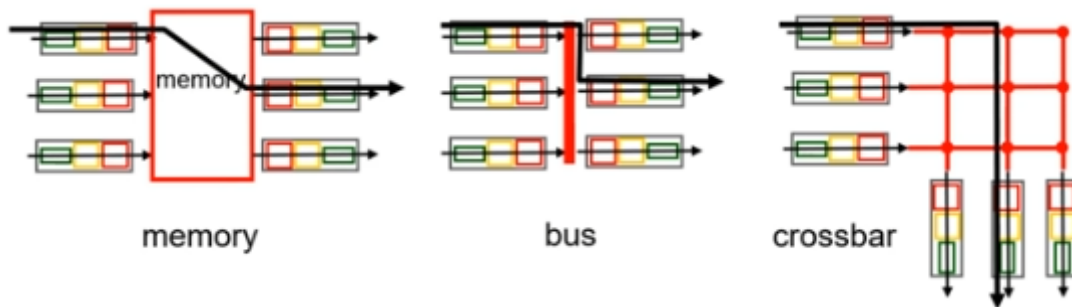
- **Packet classification**: map packets to flows
- **Buffer management**: decide when and which packet to drop
- **Scheduler**: decide when and which packet to transmit
 - Chooses among queued packets for transmission
 - Select packets to **drop** when buffer saturates

- scheduler
 - 谁优先级高给谁发



- 轮询，每个flow发一个
- 带权公平队列，权重高的多给发店

交换结构

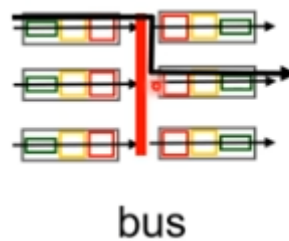


- 基于内存的



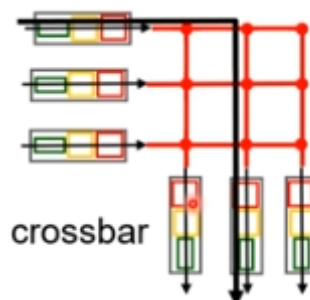
从input copy 到memory，从memory copy到output

- 基于总线的



比第一个快点，但是有总线的争用问题

-



所有的都能全速

分组交换的两种方式 Packet switching

我们之前学过链路交换circuit switching和分组交换 Packet switching

分组交换也有两种实现方式

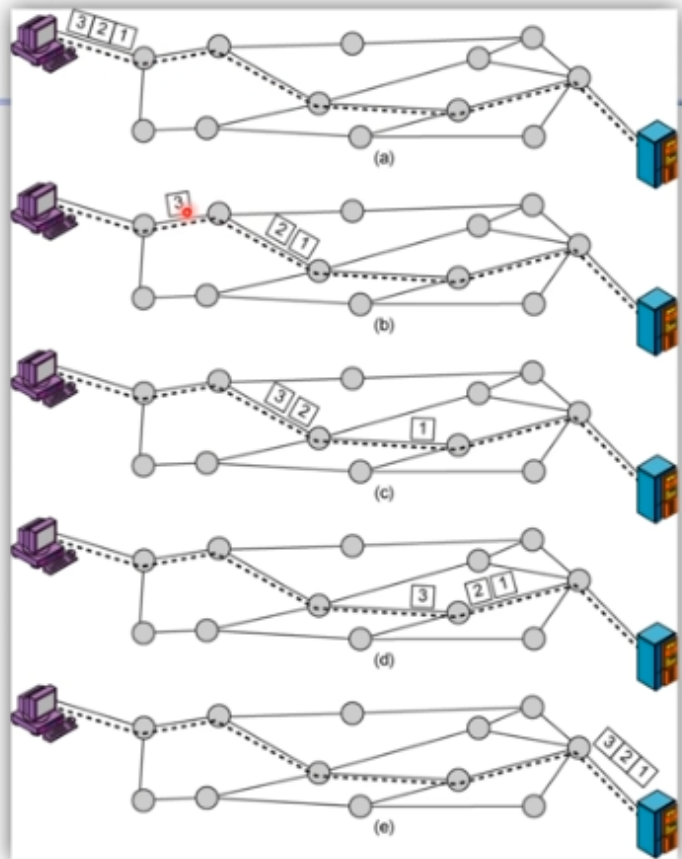
- 数据报分组交换 datagram networks--现在比较流行的IP网络
- 虚电路分组交换 virtual circuit networks--曾经兴起的ATM等网络

底层是分组交换，但是模拟链路交换，一条条的flow

虚电路走的路径唯一，按顺序到达



Routing in Virtual Circuit





Routing in Datagram Nets



数据报走的路线可以不一样，到达的顺序也可以不一样，接收方接收到之后排列顺序交给应用层，更加灵活

虚电路发送数据之前要发送方接收方 眉来眼去建立一条路径，转发表是根据VCid而不是IP地址，VCid是每一个路由器给这条路径分配的，会分配两个，一个是进来的，一个是出去的，这个VCid只对这个路由器有效