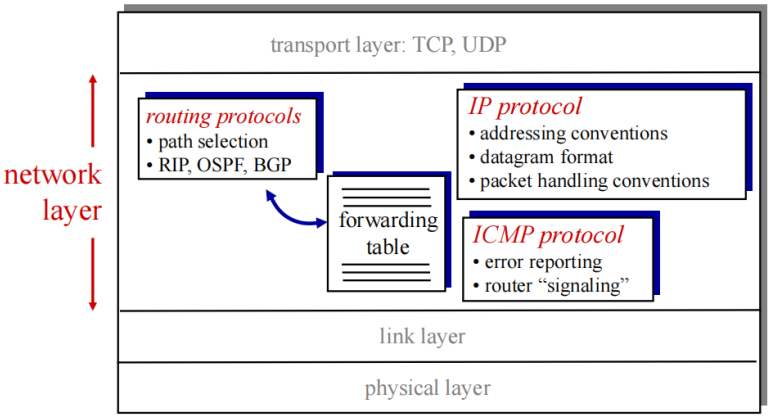
Ch4 Network Layer

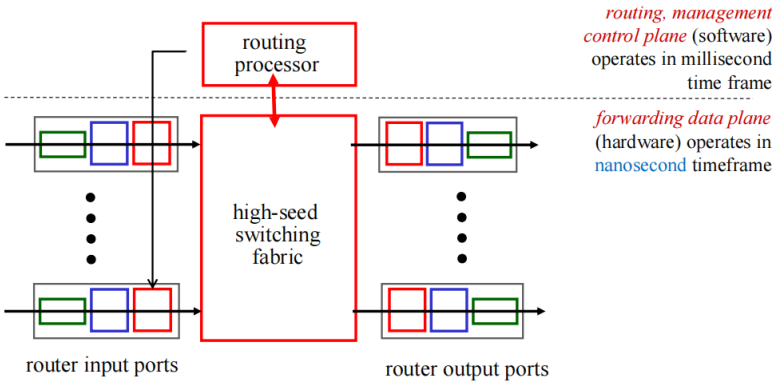
- Overview



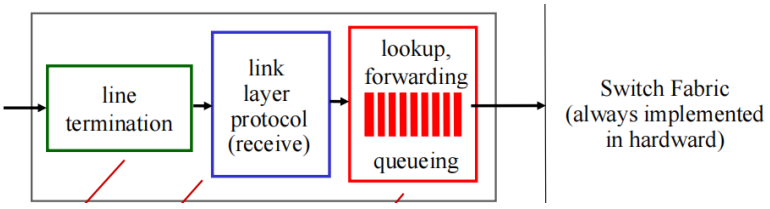
  - Def
    - host to host
    - network layer protocol 在host，router都存在
    - Two key network-core functions
      - forwarding：local 从router的input link找到合适output link
        - data plane
      - routing：global 确定src-dst path
        - routing algorithm
        - control plane
  - data plane
    - local, per-router function, hardware
    - input port -> output port
    - forwarding function
  - control plane
    - network-wide logic, software
    - source host -> dst host
    - apporaches:
      - *traditional routing algorithms:* implemented in routers
      - *software-defined networking (SDN)*: implemented in (remote) servers, 由 remote controller与每个路由器交互（control agents CAs）
  - Network service model
    - individual datagrams
      - 保证送达
      - 保证时间

- a flow of datagrams
  - 保证按顺序
  - 保证最小bandwidth

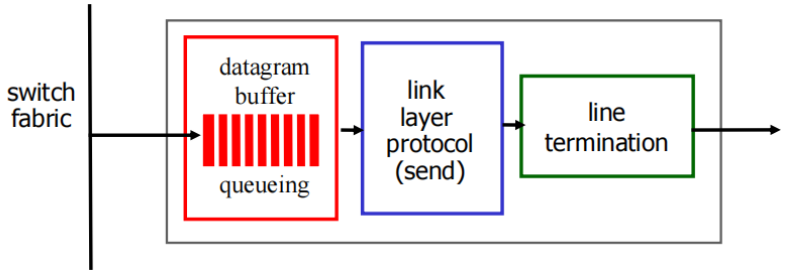
- router



  - input port

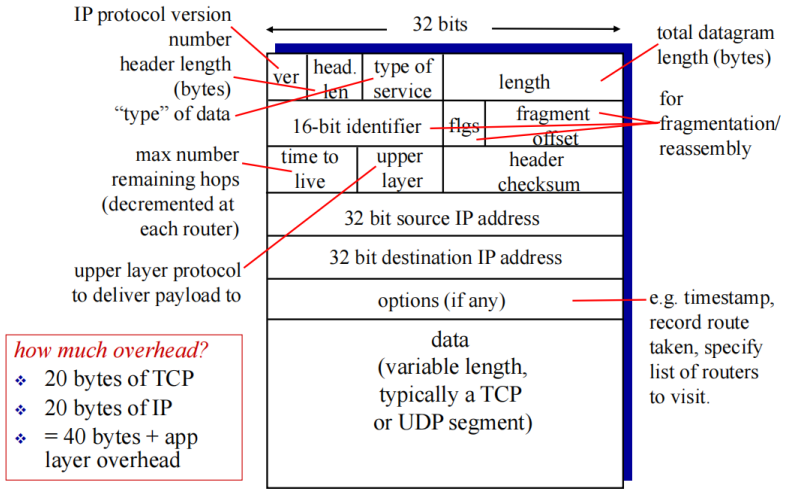


    - line termination: 接收bits
    - decentralized switching
      - 查看header，using forwarding table 看output port
      - Longest prefix matching(题)
  - switch fabrics
    - switching rate: 数据包从输入到输出的速率
    - **Switch via memory**：不能两个一起，interrupt
    - **Switch via bus**: Broadcast; label， 一次只能传一个
    - **Crossbar**: Multiple packets in parallel, Non-blocking
  - output port



    - buffering
    - scheduling discipline: Priority scheduling
  - queue
    - input port queue

        - if $R_{\text{switch}} > NR_{\text{line}}$, input port queue 可忽略
    - output port queue
        - queuing (delay) and loss due to output port buffer overflow!
        - $\text{buffer} = \frac{RTT \cdot C}{\sqrt{(N)}}$ with N flow
            - C capacity(bps)
    - scheduling: choose next packet to send on link
        - FIFO--discard policy
            - *tail drop:* drop arriving packet
            - *priority:* drop/remove on priority basis
            - *random:* drop/remove randomly
        - priority
            - *priority queuing*: high priority queue & low priority queue
        - *Round Robin (RR) scheduling*: 多类别，循环扫描队列，从每个类循环发送完整数据包
        - Weighted Fair Queuing (WFQ)：每个循环都给每个类加权
- IP: Internet Protocol
    - datagram format: 20 bytes header



    - fragmentation
        - max transmission unit (MTU) 超过就要分片
        - 到dst时候组合
        - eg 4000 B, MTU=1500B 题
            - len=1500(1480+20 header) ID=x fragflag=1 offset=0
            - len=1500(1480+20 header) ID=x fragflag=1 offset=185(1480/8) 185个bytes
            - len=1040 ID=x fragflag=0 offset=370
    - IPv4 addressing
        - IP: 32-bit identifier for interface of hosts and routers
        - routers 多个int

- host 1 or 2 (wired Ethernet, wireless 802.11)
- subnet: 可以在没有router介入的情况下reach each other
- 寻址 CIDR
  - **Classless Inter Domain Routing**
  - ISP block --- ICANN: Internet Corporation for Assigned Names and Numbers

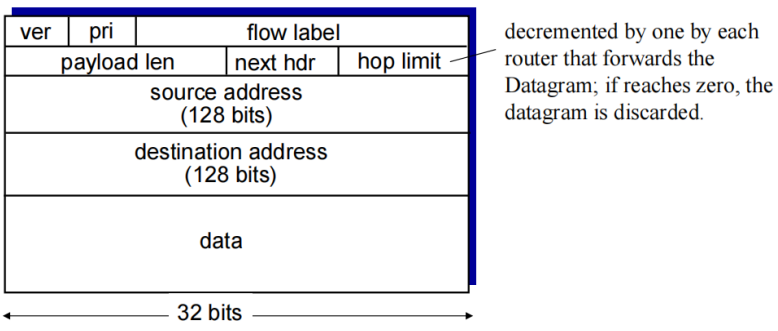
    Block of addresses: 166.4.20.128/25. Subnet 1: at least 62 interfaces; Subnet 2: at least 30 interfaces; Subnet 3: at least 28 interfaces

    This block of IP addresses can be written as
    10100110 00000100 00010100 10000000

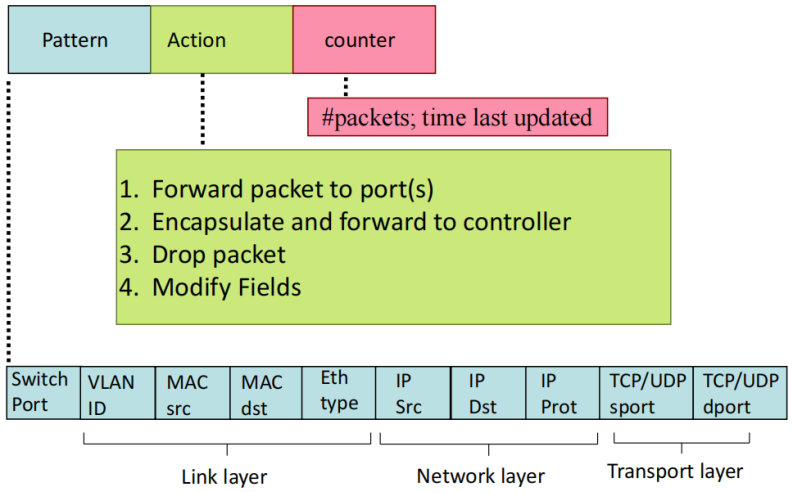
    Subnet 1: since $2^6 = 64 > 62$, we can assign the following block
    10100110 00000100 00010100 10000000
    which can be represented as 166.4.20.128/26.

    Subnet 2: since $2^5 = 32 > 30$, we can assign the following block
    10100110 00000100 00010100 11000000
    which can be represented as 166.4.20.192/27.

    Subnet 3: since $2^5 = 32 > 38$, we can assign the following block
    10100110 00000100 00010100 11100000
    which can be represented as 166.4.20.224/27.
  - route aggregation
  - How does a host get an IP?---**DHCP: Dynamic Host Configuration Protocol**
    - 自动分配IP地址及其他网络配置信息
    - DHCP Discover: client首次接入网络，广播dhcp discover，请求ip
    - DHCP Offer：分配ip，并广播dhcp offer（不知道给谁）
    - DHCP Request：client选择一个，并向选定的dhcp server 发送request表示接收
    - DHCP Acknowledgment：server->client
- **network address translation NAT**
  - 为LAN保留IP地址块，对外隐藏地址
  - 离开router的datagram都有一样的IP不同的port——可以在不更改本地addr时候更改ISP
  - 实现
    - 输出datagram：替换src IP&port为NAT IP&port
    - remember: NAT translation table src-NAT <IP&port>
- IPv6: 128-bit



    - header 40Bytes
    - 没有checksum--UDP/TCP中有，没必要，耗时
    - Tunneling解决IPv4与IPv6共存问题，在6外面套一层4，如果有router只能4
- Generalized Forward and SDN
    - intro: 每个router都包含一个flow table（control plane），定义了router的match+action
    - OpenFlow



        - 防火墙 action=drop match=IP+TCP/UDP port
        - router: action=forward, match=IP(LDP)
        - switch: action=forward, match=dst.MAC
        - NAT: action=rewrite, match=IP+port