

---

## 计算机网络总结笔记（下）

---

这部分包括网络层、链路层、物理层（未完成）和无线网络。最后给出一个总体概括和中英对照表。

### -----D.网络层-----

网络层的功能就是将分组从一台发送主机送到接收主机。需要三种功能：**转发**、**选路**和**建立连接**。而网络层应当提供的服务包括：确保交付，即数据可以到达目的地；确保交付具有时延上界，即数据交付可以确保不超过某个时延值。如果要传输分组，网络层还必须具有：有序分组交付，确保最小带宽，确保最大延时抖动，安全性服务。因特网的网络层协议 **IP** 是一种**尽力而为服务**。其它类型的网络层服务例如异步传输模型，即 **ATM**，比较重要的是恒定比特率 **ATM(CBR)**和可用比特率 **ATM(ABR)**。网络层也提供面向连接和无连接的服务，且在一个体系中只提供一种。提供前者的称作**虚电路(VC)网络**，提供后者的称作**数据报网络**。**Internet** 网络层是数据报网络。

一条 **VC** 组成为：**源和目的之间的路径**（一系列链路和路由），沿路径上每段链路的号码(称作**VC 号**)，沿每个路由的**转发表项**。一个属于 **VC** 的分组将在首部中携带一个 **VC 号**，之后每到达一个路由，路由都用转发表项中的索引值替换 **VC 号**，从而将分组发到下一条链路上。端系统向网络发送的虚电路启动和终止报文以及路由之间传递的 **VC** 建立报文称作信令报文，对应信令协议。

在数据报网络中，报文只存储目的地的 **IP 地址**，而在经过一个路由的时候，路由通过研究该目的地址在自己的分组表的哪个范围，将该分组发往相应的链路。为此路由将存储一个**前缀转发表**，并用目的地址和前缀表进行**最长前缀匹配**，详细内容看课本。数据报网络中，转发表还可以实时更新。

### -----路由工作原理-----

路由包括**输入端口**，**交换结构**，**输出端口**和**选路处理器**四个部分。

输入端口首先有**线路端接**功能，随后是**数据链路处理**，这两部分实现了通向路由的各个物理链路相关的物理层和数据链路层，之后的查找转发排队模块将运用转发表判断分组将发往哪条输出链路，将查表成功后的分组排队，之后队列中的分组交给路由的交换结构。

交换结构是路由的**核心**，三种基本的交换技术是内存交换，总线交换和互连网络交换，最后一种也是现在常用的方法。

输出端口 结构和输入端正好相反，交换后的分组进入缓存排队，之后通过数据链路层处理发往线路端。因为输入输出都有队列占用路由器缓存，如果队列太长就会耗尽缓存出现**丢包**。从两方面解决这个问题：**加大缓存**和**加大速度**。定义交换结构速率为交换结构移动分组的速率，那么，对于一个 **n** 输入 **n** 输出的路由，交换结构速率必须至少是输入速率的 **n** 倍才能够规避拥塞丢包。对于缓存容量 **B**，公式是  $B = RTT * C / \sqrt{N}$ ，其中 **C** 是链路容量，**N** 是较大 **TCP** 流流量。同样，输出队列的分组调度程序需要考虑队列中分组的传送规则，从而提供服务质量保证。这将在后面章节中讨论。如果缓存不够，可以选择丢弃新进入的分组（弃尾策略）或者删除已有分组。而在缓存满之前向发送方提供一个拥塞信号的做法称作**主动队列管理(AQM)**，**随机早期检测(RED)**是常用的 **AQM**。

如果是输入队列阻塞称作线路前部(**HOL**)阻塞。

### -----网络层-----

---

网络层分组也就是数据包的格式包括：版本号(4),首部长(4), 服务类型(8, 简称 TOS, 规定不同的数据报类型), 数据报长度(16)组成第一个 32bit, 标识、标志、片偏移组成第二个 32bit, 寿命(8, Time-to-live, TTL, 每经过一个路由该值减 1, TTL 为 0 时该数据包丢弃), 协议(8, 指定接收方的上层协议), 首部校验和组成第三个 32bit, 之后是各 32bit 的源和目标 IP, 选项内容和数据。

因为链路层协议不一定能承载网络层分组, 如果将链路层帧可以承载的最大数据量称作最大传输单元 MTU, 为了传输大的 IP 分组, 需要进行 IP 数据报分片。发送方负责给数据报加上标识, 路由器在需要拆分一个大数据报时, 为拆分出来的每个片加上片偏移和标志, 如果该片是最后一片标志为 0, 否则为 1。接收方收到一组带有同一标识的片后, 按照片偏移组装。

#### -----IPV4-----

IP 是 32 位二进制并每八位写成一个十进制, 最大的 IP 就是 255.255.255.255, Internet 中每台主机和路由必须由全球唯一的 IP。此外, 如果有若干主机处在一个没有路由器的网络中互联, 这时候加入一条连向路由器端口的链路就能形成一个子网。比如子网中三台机器是 223.1.1.1, 223.1.1.2, 223.1.1.3, 路由端口是 223.1.1.4 那么 IP 编址为该子网分配的地址就是 223.1.1.0/24, 表示 IP 的前 24 字节定义了子网地址, 这种/24 的记法称作子网掩码。Internet 的地址分配策略称作无类别域间选路(CIDR), 首先, 组织从 ISP 处获得一个 IP 地址块, 最顶级的 IP 地址分配商就是 ICANN。随后, 组织一般通过动态主机配置协议 DHCP 为组织网络中的主机和路由端口分配 IP。DHCP 是一种即插即用协议, 下面简单讨论一下 DHCP 协议的步骤。

新加入网络的主机首先要发现一个与之交互的 DHCP 服务器, 通过在 UDP 分组中向 67 端口发送 DHCP 发现报文, 使用目的地址 255.255.255.255 和源地址 0.0.0.0, 这两个特殊 IP 分别代表向全网络广播报文和本主机地址。这样这个报文就发给了全网络。DHCP 服务器收到一个发现报文时, 用一个 DHCP 提供报文响应, 目的地址同样是 255.255.255.255 全网广播, 该提供报文中包括发现报文的事物 ID, 推荐 IP 和网络掩码以及 IP 地址租用期。客户机收到后从网络的服务器中选择一个并用 DHCP 请求报文对响应服务器响应, 相应服务器用 DHCP ACK 报文确认信息, 这样客户机终于得到了一个 IP。

NAT 技术, 即网络地址转换技术是一种广泛应用的简单地址分配法。NAT 服务器对于外网相当于一个设备, 拥有一个单一 IP, 他接收所有内网数据, 并用唯一 IP 发送, 收到的数据通过路由上的 NAT 转换表发往内网特定设备。这可以有效缓解地址分配压力。但是 NAT 因为不能让外网直接到达内网设备, 由 UPnP 协议提供 NAT 穿越。该协议能让应用程序在专用 IP、专用端口和公共 IP、公共端口(也就是 NAT 路由的 IP 和端口)间建立映射, 应用就可以将公共 IP、公共端口用来对外宣告, 而将外部数据通过 NAT 接收。就像 NAT 上开了一个洞口, 穿过洞口的信息都被该程序接收, 而外界认为这个洞口就是该程序。这就是 NAT 穿越技术。

#### -----IPV6-----

数据报格式: 版本号(4), 流量类型(8), 流标签(24)组成第一个 32bit, 有效载荷长度(16), 下一首部(8), 跳限制(8)组成第二个 32bit, 之后是 128bit 的源地址和目的地址, 最后是数据段。其中, 流量类型类似 IPV4 的 TOS, 表明流类型, 流标签用于标识一个数据报的流。有效载荷长度表示该报文后面的数据长, 下一个首部字段指明数据报内容交付给什么运输层协议, 类似 IPV4 中的协议字段, 跳限制类似于 TTL。

其不同之处在于, 首先不允许在路由器分片组装, 分片工作都在端系统完成。路由如果收到大于链路容量的数据报将返回过大 ICMP 报文。此外因为运输层有首部校验, IPV6 取消了首部校验和, 选项也被移到了对应的运输层协议中。

## -----选路算法-----

因为主机肯定与一台路由器直接相连，称作默认路由或第一跳路由或默认网关，从源路由到目的路由之间我们总希望得到一条最短路。算法有**全局选路算法和分布式选路算法**。选路算法如果按照算法的状态来分可以分为**静态和动态选路算法**。按照负载敏感还是负载迟钝分为**负载敏感算法和迟钝算法**。这些概念在后面说明。

链路状态选路算法 LS 中，网络拓扑和所有链路的费用已知，这样就可以用图论中的 Prim 和 Dijkstra 算法计算最短路

距离向量选路算法 DV 则是迭代异步分布式算法。详细算法看课本。相比起来，**LS 算法收敛速度快，健壮性高，但是 DV 算法的报文复杂度低**。

在实践中，因为路由网络规模巨大，而且有组织网络管理者期望管理自治，一般将路由器组织到自治系统(AS)中。一个 AS 中的路由都执行同一种选路算法，称作自治系统内部选路协议。AS 中还有数台路由负责向 AS 外部转发分组，称作网关路由。对于网关路由来说，和他相连的外部 AS 路由当然也是网关路由。为了能够将数据转发到正确的 AS，一个网关路由必须知道和他相连的每个网关路由可以到达的路由信息，并且要把这个信息告知自己 AS 中所有路由以便让他们发送信息到对应路由。这两项任务通过自治系统间选路协议完成。

下面通过研究 Internet 的选路规则详细说明。

AS 内部选路协议又称内部网关协议，常用的是**选路信息协议 RIP 和开放最短路径优先 OSPF 的 IS-IS 协议**。AS 间选路协议则是 BGP 协议。本部分内容将在以后补充。

## -----E.链路层和局域网-----

链路层的任务是将网络层数据报通过路径中的单段链路**节点到节点传送**。其服务包括**成帧(将网络层数据报封装成帧)**，链路接入，可靠交付，流量控制，差错检测，差错纠正，半双工和全双工。链路层协议建立在**节点对相邻节点**的基础上。其功能通常是在网络适配器或称网卡(NIC)中实现的。

链路层提供比特级差错检测和纠错，手段包括采用一位奇偶校验位、或者二维奇偶校验，也就是把数据分成 i 行 j 列对行列分别求奇偶校验值，二维奇偶校验有自动纠正一位错误的能力。或者是用在网络层中用的校验和的方法。接收方检测和纠正差错的能力称作前向纠错(FEC)

现在网络中采用**循环冗余检测(CRC)**编码，也称作多项式编码。

## -----广播链路-----

网络链路有两种类型，**点对点链路和广播链路**。广播链路需要涉及到多路访问问题，也就是图和协调多个发送和接受节点对一个共享广播信道的访问。如果两个节点同时传输帧并且帧发生**碰撞**，两个节点都无法收到这个帧。网络中对应多路访问协议。这种协议的实现可以分为三类：**信道划分协议，随机接入协议和轮流协议**。

信道划分协议：**时分多路复用 TDM 和频分多路复用 FDM** 是两种信道划分协议。TDM 将时间划分为**时间帧**，并将时间帧划分为 N 个**时隙(slot)**，然后把时隙分配给节点，随后按照分配好的**顺序循环**。一般一个时隙可以传输一个分组。这样可以有效避免碰撞。但是也可能造成浪费（比如只有一个节点要发信息，他每次都要等其他节点耗完他们的时隙）。FDM 则把信道带宽划分**成 N 段**，每个节点分配到信道带宽的一部分，优缺点和 TDM 一样。第三种信道划分协议就是码分多址 CDMA 协议。通过巧妙为**每个节点分配一种编码**实现不同节点能同时发送并不碰撞。细节将在以后的章节中讨论。

随机接入协议：随机接入协议感觉和拥塞机制比较像。当节点经受碰撞后，他将选择一个**随机时延**，之后重发该分组，重复此过程直到分组发送成功。最常用的是**ALOHA 协议**和**载波侦听多路访问 CSMA 协议**。

时隙 ALOHA 协议中假定所有帧都是  $L$  bit 长，时隙长为  $L/R$ ， $R$  为信道速率，也就是说**一个时隙传输一帧**，节点只在时隙起点传输帧，节点同步，节点能在时隙结束前得知碰撞。时隙 ALOHA 中，如果发生碰撞，节点将以  $p$  概率在下一时隙中重传，以  $(1-p)$  概率在下一时隙中再进行一次概率运算，直到传输成功。这个简单协议下最大效率大约仅为 37% 纯 ALOHA 是非时隙分散协议，其最大效率仅为时隙 ALOHA 的一半。

载波侦听多路访问协议 CSMA 协议中，载波侦听要求节点传输之前**先侦听信道**，如果发现节点在信道上传输，就等待随机时间再侦听信道，**碰撞检测**要求节点在**传输时侦听信道**，如果检测到另一节点在传输干扰帧。就停止传输，并由协议规定其下次尝试时间。因为侦听信道可能有传输延迟，也就是说某一个节点已经在传输，但是另一个节点侦听到信道是空的，这样还会发生碰撞，碰撞检测 CSMA 中，这两个节点检测到碰撞后，都将放弃本次传输。

轮流协议中比较重要的是**轮询协议**和**令牌传递**协议。轮询协议中将一个节点规定为**主节点**，由它轮流给每个节点发送报文 指定传输速率等。令牌传递协议中有一个小的目的帧称作令牌，令牌在节点间以某种固定次序轮换，持有令牌的节点可以以最大速率发送一些帧。

局域网(LAN)有两种流行的技术，一种是基于随机接入的**以太网 LAN**(或称 802.3 LAN)，一种是**令牌传递技术的 LAN**，包括令牌环(或称 IEEE802.5)和光纤式分布数据接口(FDDI)

-----链路层编址-----

LAN 地址也被称作物理地址或者 MAC 地址，MAC 即媒体访问控制地址。一般 MAC 地址为 6 字节，并且任意适配器的**MAC 地址是唯一的**。所以现在很多注册软件都把注册码绑定到 MAC。发送适配器将目的 MAC 插入到帧中，LAN 中所有适配器将收到该帧并和自己比较，不一样就丢弃，一样的话就取出封装的数据报，并中断其父节点。如果需要广播一帧，就将 MAC 设置为全 1。

为了转换网络层地址如 IP 和链路层地址如 MAC，产生了**地址解析协议(ARP)**，节点上的 ARP 模块取任意一个 LAN 中的 IP 作为输入，输出对应的 MAC。和 DNS 很类似。ARP 模块在 RAM 中存储一个**ARP 表**，存储了 IP 到 MAC 的映射关系和一个 TTL 值，TTL 时间到将删除这个映射。而映射的增加过程是这样的：一个节点向适配器传递一个 ARP 分组，要求适配器用 MAC 广播发送。适配器封装后将分组传入子网中，每个节点都检查他的 IP 是否和分组中的目的 IP 一致，如果一致就发回响应 ARP 分组，这样发送节点就能够更新他的 ARP 表了。注意**ARP 分组是一个广播帧**，但是**响应 ARP 是一个标准帧**。如果想要跨路由器从一个子网发数据到另一个子网，则要首先发送到路由器，由路由器对照内部的转发表转发。

-----以太网-----

之前说过以太网是现在流行的有线局域网技术。以太网帧包括前同步码(8 字节)，目的地址(MAC 地址 6 字节)，源地址(6)，类型(指定网络层协议，2 字节)，数据(承载 IP 数据报，最大传输单元 MTU 是 1500 字节)，最后是循环冗余检测的 4 字节。此外以太网向网络层提供**无连接不可靠服务**。

以太网使用的多路访问协议是**CSMA/CD** 协议。它允许适配器在**任意时刻开始传输(无时隙)**，**载波侦听**，**碰撞检测**，尝试重传前有**等待随机时间**。其工作流程则是，首先适配器从网络层得到数据报并准备一个帧，如果适配器侦听到信道空闲(96bit 时间内无**信号能量**进入信道)，就发送这个帧，如果侦听到其它信号能量，就一直等到信号能量消失并再等待 96bit 之后发送。发送过程中同样要侦听。如果侦听到其它信号能量就停止传输转而发送一个**48bit 的阻塞信号**。



---

之后，适配器进入**指数后退**阶段，即当该帧面临第  $n$  次碰撞时，适配器随机从  $0 \sim 2(m)-1$  中选一个值  $K(m=\min(10,n))$ ，之后适配器等待  $K*512$  比特时间。

**以太网效率** $=1/(1+5*d[\text{prop}]/d[\text{trans}])$ ，其中  $d[\text{prop}]$  为信号能量的最大传递时间， $d[\text{trans}]$  为帧的最大传输时间。

以太网标准很多，其统一命名规则是传输速率数字+BASE+物理媒体字母。如 10GBASE-T 代表标准速率为 10Gbps 的双绞铜线以太网。

#### -----交换机-----

现代以太网是**星形拓扑**的。每个节点都和中心交换机相连。交换机的任务是**接收入链路层帧并转发到出链路**。交换机对于节点是**透明**的。也就是说**节点寻址不通过交换机**。

交换机有**转发**和**过滤**功能，借助于内部的交换机表完成交换机表包含节点的 MAC 地址，连接该节点的接口，和这个映射的保存时间。其工作原理和路由很像，也是**自学习**和**即插即用**的。交换机来连接以太网中的节点，首要好处就是**消除了碰撞**，节点需要传输帧的时候，就像路由一样，交换机会把帧放到**输出缓存队列**中。交换机还能连接异质链路，让不同链路共同工作，交换机也易于管理。

和路由器相比，交换机**最高处理至链路层数据**，但是路由可能要处理网络层数据，但路由的保护措施可以保证网络不崩溃。此外，交换机是自学习的，而路由需要设置额外的 IP 地址和 MAC，**交换机能直通交换，但路由需要通过路由转发**。

#### -----PPP 协议-----

广播链路一节开头说，链路层协议主要是广播链路协议和点对点协议。**点对点协议即 PPP**，也是主机拨号所选择的协议。其内容看书。

#### -----F. 无线网络 -----

无线网络由无线主机，无线路由和基站组成。基站负责在主机间交换数据，实例有蜂窝网络中的蜂窝塔(cell tower)，无线 LAN 的接入点。与基站关联的主机被称作基础设施模式运行，另一种模式就是自组织网络模式。我们讨论前者。无线链路和网路特点包括路径损耗产生衰减，其它信号源干扰，多径传播，这些特征使得无线网络更容易发生比特错误，衡量的值是比特差错率 (BER)，对于无线主机，接收到的信号和噪声的比例称作信噪比 (SNR)，一中物理层调制方案下，SNR 越高，BER 越低。

在广播链路一节中，我们说过三种访问协议，信道划分协议中包括的码分多址 (CDMA) 协议。CDMA 中，每个待发送 bit 都乘以一个信号 bit 进行编码，这个信号的变化速率，也就是所谓码片速率比数据的速率快很多。具体算法课本上很详细。

#### -----WIFI(802.11 无线 LAN)-----

WIFI 基本构建模块是基本服务集 (BSS)，通常包括多个无线站点和一个称作接入点 (AP) 的中央基站。配备 AP 的无线 LAN 称作基础设施无线 LAN。每个 AP 都被分配了一个服务集标识符 (SSID) 和一个信道号。设备进入 WIFI 丛的时候需要选择一个 AP 进行关联。标准要求 AP 周期发送信标帧，包含该 AP 的 SSID 和 MAC 地址，对于主机来说，扫描信道监听信标帧称作被动扫描，也可以通过广播一个探测帧进行所谓主动扫描。WIFI 的高级特色包括速率适应和功率管理。其他 IEEE 802 协议包括蓝牙和 WiMAX。

802.11 协议也有类似以太网的随机访问协议，称带有碰撞避免的 CSMA，即 CSMA/CA 协议。其特点是使用碰撞避免而非碰撞检测，并采用 ARQ 机制。碰撞避免类似于 TCP 的 ACK 机制。

---

发送方侦听到信道空闲后经过一个分布式帧间间隔(DIFS)后发送该帧,接收方经过短帧间间隔(SIFS)后发回确认帧,如果发送方没有在规定时间内收到确认就跟以太网一样用回退机制重发。多次尝试失败后发送方就抛弃这个帧。

但是在侦听信道方面有个大问题。因为无线设备覆盖区域有限,如果两个发送方都不在彼此的覆盖范围内,但是都能向一台接收机发送数据,那么他们还是会在自己不知道的情形下在接收方处发生碰撞。为了处理这个所谓的隐藏终端问题,协议包括了 RTS 和 CTS 缓解这个问题。发送方发送数据之前发送一个请求发送 RTS,接收方发回 CTS 通知可以传输,另外要求其他发送方推迟访问。

下面看 IEEE 802.11 帧规范。每一帧包括(单位:字节)帧控制(2),持续期(2),地址 123(各 6),序号控制(2),地址 4(6),有效载荷(0~2312),CRC(4)。其中帧控制的子字段包括(单位:bit)协议版本(2),类型(2),子类(4),到 AP 标识、从 AP 标识,更多标识,功率管理,更多数据, WEP, Rsvd(各 1)。前三个地址分别是接收方 MAC,发送方 MAC,子网路由 MAC。最后,蜂窝网络和 GSM 等略。



分割线-----

现在来整体看一下网络的运作过程,也算是总结之上的总结吧。

互联网是主机和链路组成的网络。互联网中所有终端都被分配了一个 IP,所有网卡都有唯一的 MAC 地址。端系统上的应用使用应用层协议形成网络请求,请求报文中包括目的服务器的域名,域名通过 DNS 解析成 IP 地址,通过套接字传递给本机的运输层,运输层的主要协议是 TCP 和 UDP,他们都能提供逻辑通信,不同的是 TCP 还提供了拥塞机制和可靠数据传输。应用层可以根据需要指定运输层协议。之后,运输层把封装好的报文交给操作系统管理的网络层。网络层提供转发和选路功能。我们首先讨论了路由器的转发和选路以及缓冲队列等功能是怎么实现的,随后指明网络层协议 IPV4 和 IPV6 的数据报结构,以及他们提供不可靠的数据传输服务,还说明了 IP 分配方面的内容,即为了子网管理方便发明的子网掩码,为了主机临时接入方便的 DHCP 协议,为了解决 IP 地址紧张,简化网络结构发明的 NAT 技术,和 P2P 程序为了 NAT 穿越采用的 UPnP 协议。在网络层部分的最后讨论的 DV 和 LS 选路算法。网络层数据报最终交给了链路层封装成帧。链路层提供节点到节点之间的数据传输。为了解决广播链路信道上的碰撞问题,我们讨论了三种多路访问协议:信道划分协议,随机接入协议和轮流协议,最终说明了 LAN 使用的以太网协议。对应网络层的内容,我们简单说明了每个适配器都有唯一 MAC 地址,以及寻址和交换机的内容,最后讨论了另一种网络链路:PPP。另外,对于换汤不换药的无线移动网络做了简单介绍。

-----中英文对照表-----

转发 forwarding 选路 routing 转发表 forwarding table 连接建立 connection setup  
异步传输模型 ATM 恒定比特率 ATM Constant Bit Rate CBR  
可用比特率 Available Bit Rate ABR 虚电路 virtual Circuit VC 数据报网络 datagram  
network 信令报文 signaling message 最长前缀匹配原则 longest Prefix matching rule

---

弃尾策略 drop-tail 随机早期检测 Random Early Detection RED 主动队列管理  
Active Queue Management 最大传输单元 Max Transmission Unit MTU 片 fragment