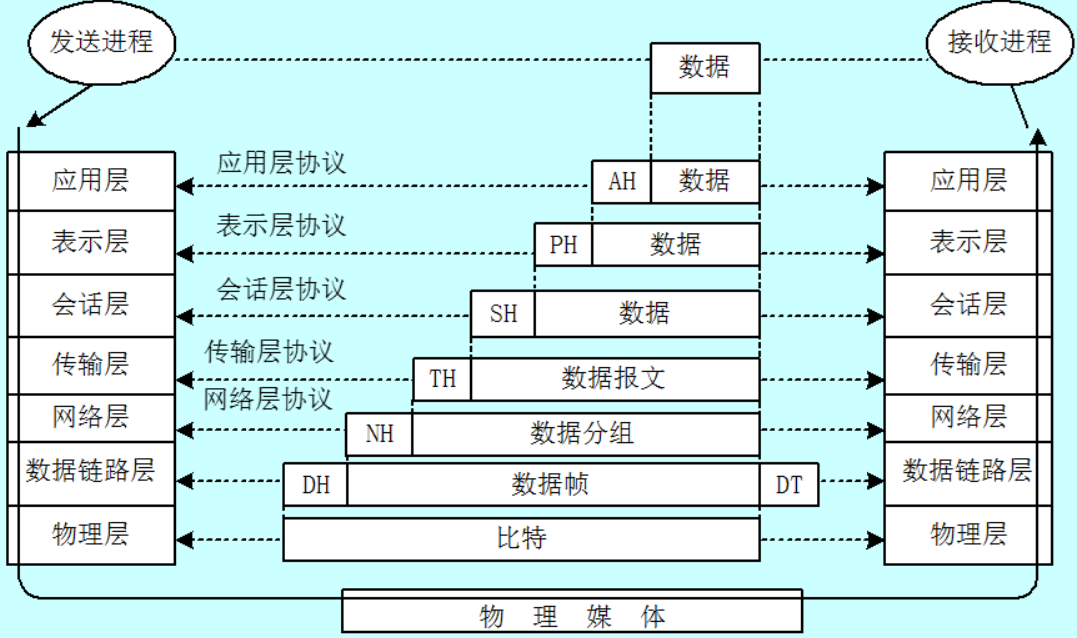
# 计算机网络体系结构



## 计算机网络概述

网络分类

局域网：Local Area Network,LAN

城域网

广域网：Wide Area Network，WAN

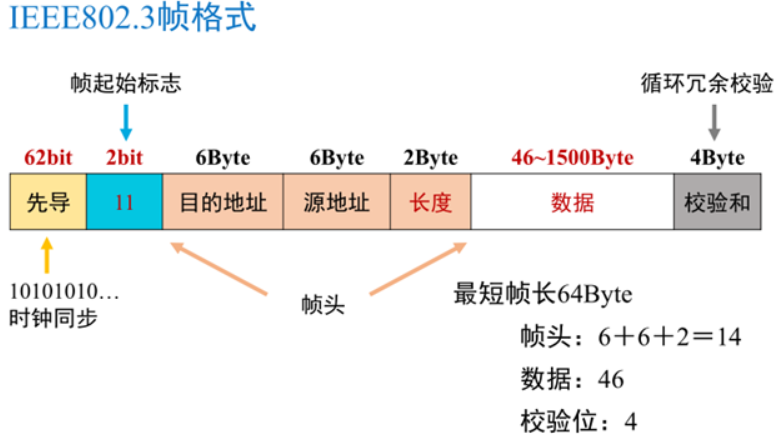
## 1.2 计算机网络体系结构

### OSI/RM网络体系结构

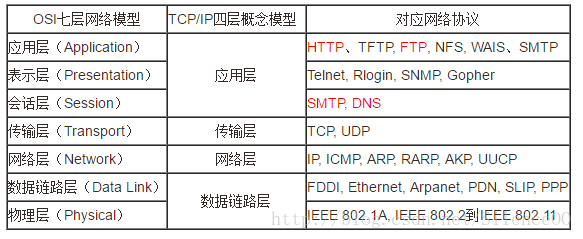
物理层：提供物理连接，包括传输介质和部分网络设备。

数据链路层

Ethernet帧结构



### TCP/IP体系结构



## 1.3 网络的基本拓扑结构

### 1.3.1 网络拓扑结构类型

网络设备（计算机，路由器等）与链路的互连构成的系统。

#### 1.3.1.1 点对点网络

Peer to Pear：两两相连，用于城域网和广域网。

端到端(End to End)：跨越多个节点。是一个逻辑链路，通信完成链路就释放，TCP是用来建立这种端到端链接的协议。

#### 1.3.1.2 广播型网络

局域网大多采用广播传播方式。

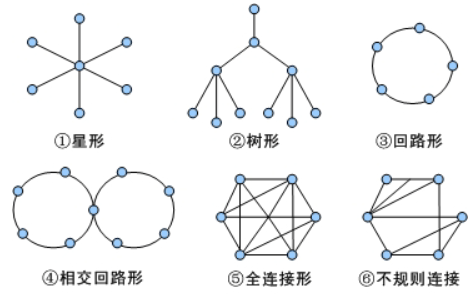
传播方式：单播，多播和组播。

信号从发送主机向共享电缆的两端传播，连接在共享电缆上的所有主机都可以接收到。注意信道同时只有一个数据帧传输，即数据帧独占整个电缆，一般采用CSMA/CD。

优点是任何两个节点之间的通信最多只需要“两跳”的距离。缺点是流量大时性能下降。点对点恰好相反。

广播型网络利用传输介质的共享性，消除网络线路的重复性建设，有重要的经济意义，因此广泛用于局域网。

拓扑结构图



### 1.3.2 点对点网络拓扑结构

常见拓扑结构

线型网，环形网，网状网等

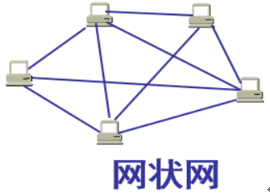
线型

由两个或多个节点串联在一条线路上。

○---○---○---○

网状网

任何两节点都有直达链路连接，通信建立过程不要信号转接。



环形网

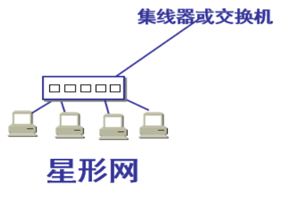


### 1.3.3 广播型网络拓扑结构

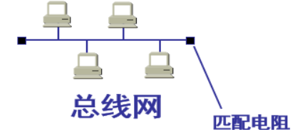
拓扑结构：总线网，星型网，蜂窝型网等。

星型网络拓扑

通过中心节点（如Hub，交换机）相连，所有数据都经过中心节点进行交换。



总线型拓扑



优点：不需要互联设备（如Hub，交换机），组网费用低。

缺点：共享同一总线带宽。。。

# 物理层与网络设备

## 2.6 传输介质

## 2.7 物理层网络设备

2.7.1 调制解调器

2.7.2 集线器

2.7.3 光线收发器

## 2.8 其他层网络设备

### 2.8.1 网卡

网络适配器(Network interface Card,NIC)，连接物理层和数据链路层。

### 2.8.2 交换机

1.网桥

数据链路层设备，连接两个同构相互独立的计算机网络。这里同构主要指网络的拓扑结构相同，网络协议相同。

网桥按MAC地址转发数据帧，没有路由能力，只能连接两个相同的网络。

网桥中每个网络端口维护一个MAC地址表，有路径学习能力。

2.交换机

以太网交换机从网桥发展而来。通常使用硬件实现过滤，学习和转发数据帧

第2层交换机

在数据链路层实现，交换机内部维护一个交换矩阵（MAC地址表）。根据MAC地址转发数据，无需源主机和信宿主机的IP。

第3层交换机

IP转成MAC，仍然通过MAC地址通信。大大减少拆包/打包工作，速度高于路由器。

第三次交换机可以隔离广播数据包。

### 2.8.3 路由器

1.网关

连接两个异构的相互独立的网络。3种网关类型：协议网关/应用网关/安全网关，不同网关工作在不同的网络层次。

早期往往将路由器称网关。

2.路由器

通过转发数据包实现网络互连。

2.路由器的主要功能

* 协议转换：数据格式转换可由硬件完成，也可由软件实现。
* 路由功能
* 网络互联

2.8.4 防火墙

# 第三章 数据链路层与局域网技术

## 3.1 数据链路层工作原理

### 3.1.1 数据链路层的功能

基本概念

链路：点到点的物理线路，中间不存在交换节点

数据链路：链路+必要规程控制。生存期为通信开始到结束止。

链路类型

点对点链路：一对一，如PPP

广播式链路：一对全部，如以太网，无线网

交换式链路：一对组

协议类型

广播链路（局域网）：IEEE802.2,IEEE802.3

点对点链路（广域网）：PPP，LAPM

## 3.2 广域网数据链路层

### 3.2.2 链路层协议

电话拨号接入

用户接入因特网一种是个人通过电话线路拨号接入互联网，不需申请IP。另外一种是专线接入（多为企业或单位），一般需申请IP。

因特网服务提供商（ISP），提供拨号入网。ISP采用路由器和专线与因特网相连。

拨号接入采用点对点的数据链路协议。

以太网的点对点协议

用于ADSL和Cable Modem因特网接入服务。

常见协议：PPP，PPoE，HDLC等。

PPoE：允许两个以太网端口连接，建立IP层的隧道协议，用于对另一个以太网主机进行虚拟拨号。提供标准的PPP服务，比如认证，授权等

## 3.3 以太网数据链路层

以太网(Ethernet)：1973年出现，应用广泛。以他为基础形成IEEE802.3标准。

### 3.3.1 IEEE 802 网络体系结构

1.早期局域网模型

局域网模型与广域网由很大的区别。由于局域网只是一个计算机通信网，早期局域网没有考虑路由选择问题，没有网络层，只有物理和链路层。

### 3.3.2 以太网介质访问控制子层

#### 1.MAC帧格式

#### 2.MAC子层工作过程

发送过程：生成MAC帧格式的数据。

接收过程：解析接收的比特流，本机MAC地址与目标地址比较，不匹配则丢弃。

#### 3.信道共享技术

静态分配：频分复用（FDM）或波分复用（WDM）。

动态分配：CSMA/CD。

#### 4.CSMA/CD工作原理

以太网，信道的动态分配采用，控制多个用户共用一条信道的协议。适合总线结构的网络。

载波监听（先听后发）：

### 3.3.3 交换式以太网技术

即采用交换机构成的网络。

## 3.5 虚拟局域网技术

### 3.5.1 工作原理

vlan基本概念

建立在以太网交换机之上,以软件方式实现的逻辑网络。不受物理位置限制，可以在同一局域网交换机，也可以在不同局域网交换机。

设计目的：为用户端口提供独立的广播域(OSI/RM第三层也提供了广播域分段功能)。

同一VLAN之中的主机可自由通信，不同VLAN之间通信必须通过路由器或3层交换机。

工作原理

Valn重新划分lan成员之间的逻辑连接关系,连接在一个交换机或处在一个ip子网的主机，他们之间通信受限。成员间寻址不再简单根据MAC地址或ip，而是根据vlan卷标中的寻址结构字段VID（vlan标记）。

### 3.5.2 vlan划分方法

基于端口

基于mac地址

# 第四章 网络层与路由技术

## 4.1 IP工作原理

4.1.1 网络层的功能与主要协议

主要功能：把IP分组从一个网络传送到另一个网络。

主要协议

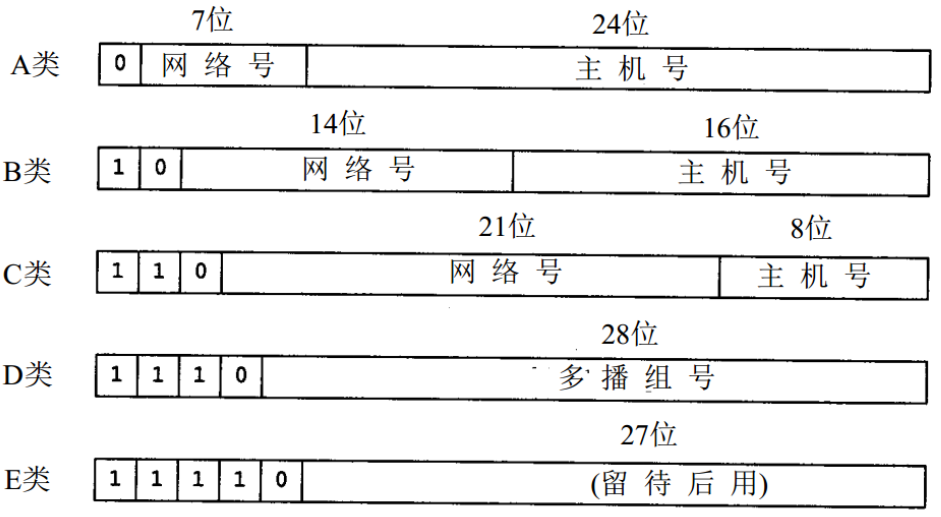
* IP:地址转换，分组处理。
* 路由选择协议：决定IP分组发送过程中由源地址到目标地址所经过的路径。如RIP/OSPF
* ARP：解决网络层地址（IP地址）与数据链路层地址（MAC地址）映射问题。只能用于具有广播能力的网络（如Ethernet）。

### 4.1.2 IP分组格式

## 4.2 IP地址分类与子网化技术

### 4.2.1 IP地址分类

IP地址=网络地址+主机地址



A类地址：0（开头）+7位（网络地址）+24位（主机地址）

B类地址：10（开头）+14位（网络地址）+16位（主机地址）

C类地址：110（开头）+21位（网络地址）+8位（主机地址）

D类 以1110开始 用于组播

E类 以11110开始 用于科研保留

地址范围

A类 从1.0.0.0 到126.255.255.255

B类 从128.0.0.0到191.255.255.255

C类 从192.0.0.0到223.255.255.255

其中127.x.x.x段地址空间是被保留的回环地址

私有地址

* A类: 10.0.0.0 ～ 10.255.255.255
* B类: 172.16.0.0 ～ 172.31.255.255 (网络号范围： 10101010 00010000 ～ 10101010 00011111)
* C类: 192.168.0.0 ～ 192.168.255.255(网络号范围： 11000000 10101000 00000000 ～ 11000000 10101000 11111111)

### 4.2.3 有类子网划分技术

1.网络的子网化

原来IP地址的主机位借出若干高位作为子网标识。

2.子网掩码

子网掩码连续为1时，对应IP地址值为子网号，连续为0时，对应主机号。

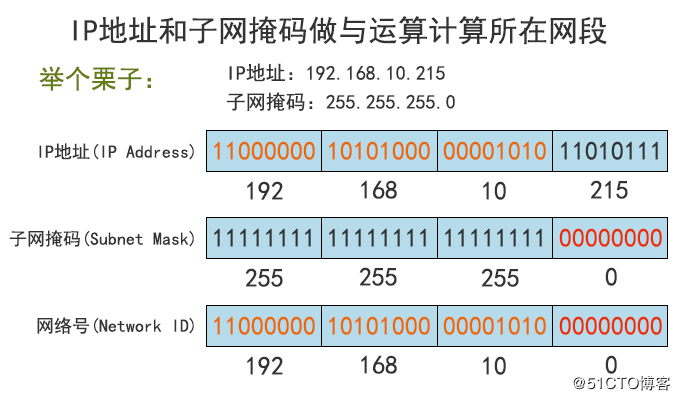
192.168.10.25

子网划分前：子网掩码 255.255.255.0

子网划分后：255.255.255.240

划分后的网络：

244：1111 0000



3.子网划分

子网划分主要解决网络号不够用，而不是解决IP不够用，因为子网划分使IP地址减少。子网之间的通信需要使用路由器。

大多数企业内部用局域网，与外部网络的互连往往采用NAT技术。

4.默认网关地址

默认网关地址是本地子网中路由器的IP地址。当放送数据的计算机发现目标地址不在本地子网络时，就将数据发送给默认网关，而不是直接向目的计算机发送。

路由器有两个网络接口和两个IP地址。

# 第五章 传输层

## 5.1 传输层基本功能

基本功能

* 应用进程之间的寻址：网络层通过IP标识主机，传输层通过端口表示一台主机的进程。
* 提供数据传输的可靠性
* 流量控制
* 堵塞控制

传输层原语

Listen,Connect,Send,Receive,Disconnect

TIME\_WAIT

产生原因

解决方案

Socket

套接字。Socket=协议类型+IP地址+端口号。

Linux内核集成了TCP/IP，提供了Socket接口。

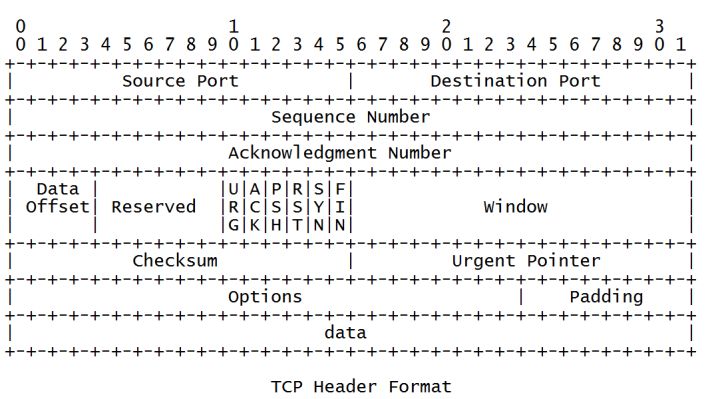
## 5.3 TCP工作原理

[Reference](https://github.com/jawil/blog/issues/14)

### 5.3.1 TCP概述

TCP是面向连接的传输层协议，提供双工可靠的服务。

### 5.3.2 TCP报文格式



一行32位 4byte，options和padding可选，所以header至少20bytes

TODO: 通过浏览器看到的request header存在哪个字段？

Header

Sequence Number 标识发送数据

Acknowledgment Number：接收端确认数，并标识期望序号

Data offset：头部长度。

Reserved：保留字段，供今后使用。

Window：窗口大小。发送方接收窗口大小。

URG：紧急比特。URG=1时表尽快传输，不要按原来的排序传输。配合Urgent Ponter使用。

PSH：PSH=1表示请求远程TCP进程立即传输给应用层，不要等缓冲区满再向上层交付。

ACK：确认比特。只有ACK=1才有意义。

SYN：SYN=1表连接请求，SYN=1，ACK=1表接受报文。SYN=0，ACK=1表连接完成。

RST：RST=1表出现严重差错，需重建连接。

FIN：FIN=1表字符串传输完成，要求释放传输连接。

Urgent Ponter：告知接收方紧急数据有多长。注意即使窗口大小为0也可以发送紧急比特。

Options

Data

应用层数据包，比如HTTP

TODO 协议大小 [Refer](http://www.ruanyifeng.com/blog/2017/06/tcp-protocol.html)

### 5.3.3 TCP连接的建立与终止

建立连接



三次握手

第一次：连接方发送连接报文SYN=1，发起连接。seq=x初始化报文序号

第二次：对方同意发回SYN=1,ACK=1(连接接收报文)。

第三次：连接方发送SYN=0,ACK=1,完成连接。

连接终止

当通信的一方没有数据要发送，使用FIN发送关闭请求，这时虽然不发送数据，但不排斥接收数据，只有对方也递交终止请求连接才会完全终止。实际也是三次握手过程。



客户端收到FIN=1才Close,

**State**

LISTEN 服务监听

ESTABLISHED 意思是建立连接。表示两台机器正在通信

CLOSE\_WAIT 对方主动关闭连接或者网络异常导致连接中断

TIME\_WAIT: 我方主动调用close()断开连接，收到对方确认后状态变为TIME\_WAIT

**TIME\_WAIT占用的资源：**少量内存（查资料大概4K）和一个fd。

### 5.3.4 数据传输的可靠性

IP分组传输过程可能出错，丢失，乱序或重复，TCP通过差错控制提供可靠性。

差错控制通过3个功能完成：checksum，确认和超时重传。

Checksum：出错丢弃数据。

确认和重传：发送方发送字节流以字节编号，接收方收到编号正确的数据，给对方确认报文。收到确认报文就可继续发送后续数据。但报文可能中途丢失，所以发送报文后就启动一个定时器，超时重传，收到确认报文取消定时器。

如何保证传输可靠性？

## 5.4 UDP工作原理

面向无连接，不保证可靠的数据传输，不具有确认重传等机制，而是依靠应用层协议处理。

无连接协议：发送前不建立连接

## 5.5 流量控制

防止网络堵塞与死锁。

### 5.5.1 滑动窗口机制

TCP采用大小可变的滑动窗口进行流量控制，以消除接收方缓存溢出。发方可连续发送一定数量字节的数据而不必等待对方确认，通过窗口字段控制上限。

滑动窗口

滑动窗口就是两台主机之间传输数据时的**缓冲区**。每台主机有接收/发送两个窗口。

TCP头部窗口字段是主机给对方设置的发送窗口上限，通信过程，接收方可根据自己资源情况动态调整窗口大小。

工作原理

TCP从应用层接收到的数据放入发送窗口，将带序号的报头加入数据包并交给IP，由IP发送到目标主机。

数据包到达接收窗口按序号放置，当接收到一个连续的数据段后，接收方向源主机发送一个ACK信号，其中带有当前窗口尺寸。

源主机接收到确认信号，将发送窗口滑到等待发送的数据区。

### 5.5.2 TCP的拥塞控制

# 第六章 应用层

## 6.4 Web服务器