

## 第一部分 计算机发展及应用

### 1) 计算机的发展

- **1946**年2月，在美国宾夕法尼亚大学诞生了世界上第一台电子计算机 **ENIAC** (Electronic Numerical Integrator And Computer)，这台计算机占地 **170** 平方米，重 **30** 吨，用了 **18000** 多个电子管，每秒能进行 **5000** 次加法运算。
- 1958年中国研制了第一台电子管计算机，速度每秒二千次。
- “银河”1983年问世，运算速度为每秒1亿次。
- 2001年，中科院计算所研制成功我国第一款通用CPU——“龙芯”芯片

### 2) 计算机的名人

- **冯·诺依曼** ——计算机之父  
1944年8月至1945年6月，美籍匈牙利科学家冯·诺依曼提出了一个全新的计算机概念，即“冯·诺依曼计算机”模型。  
主要贡献：确立了现代计算机的基本结构。
- **艾兰·图灵** ——人工智能之父，**图灵机**  
“诺贝尔奖”是世界上最负盛名的奖项，但仅用于奖励那些在物理、化学、文学、医学、经济学与促进世界和平等方面做出开拓性重大贡献的人士。“**图灵奖**”则是**计算机科学领域的最高奖项**，有“**计算机界诺贝尔奖**”之称。设立这个大奖，既是为了促进计算机科学的进一步发展，也是为了纪念一位天才数学家、计算机科学的奠基人艾兰·图灵。  
图灵是走在时代前面的天才，**在计算机远未问世之前，他构造出一台完全属于想象中的“计算机”，数学家们把它称为“图灵机”，而冯·诺依曼相当于把图灵虚构的“图灵机”进行了实现。（图灵机不是一台真正的计算机）**
- 摩尔 ——摩尔定律  
1965年，英特尔联合创始人戈登·摩尔提出以自己名字命名的“摩尔定律”，意指集成电路上可容纳的元器件的数量每隔18至24个月就会增加一倍，性能也将提升一倍。  
(摩尔定律过去是每5年增长10倍，每10年增长100倍。而如今，摩尔定律每年只能增长几个百分点，每10年可能只有2倍。因此，摩尔定律结束了。)
- 比尔·盖茨  
比尔·盖茨 (Bill Gates)，全名威廉·亨利·盖茨三世，简称比尔或盖茨。1955年10月28日出生于美国华盛顿州西雅图，企业家、软件工程师、慈善家、微软公司创始人。曾任微软董事长、CEO和首席软件设计师。  
微软公司在个人计算机和商业软件、服务和互联网技术方面都是全球范围内的领导者。我们所用的 windows 操作系统系列，office 等应用软件都是他们的产品。

### 3) 计算机的特点

- 速度快——计算机的运行速度以 MIPS 衡量。（MIPS——每秒可执行的百万条指令）
- 计算精度高——主要取决于计算机的字长
- 记忆能力强——有存储器
- 可靠的逻辑判断功能——能进行逻辑运算
- 有自动控制能力——具有程序控制下的自动执行能力

### 4) 计算机语言的发展

#### • 第一代：机器语言

由 0 和 1 组成，能让计算机直接接受的代码称为机器指令。

所谓机器语言是指计算机指令的集合

机器语言的缺点：难学、难写、难记、难检查、难修改

#### • 第二代：汇编语言

由于机器语言难以记忆，汇编语言使用了一些助记符来表示每一条机器指令。但汇编语言要通过汇编程序翻译成机器语言才能执行

汇编语言的缺点：只适应于特定的计算机

#### • 第三代：高级语言

第一种高级语言是 fortran

比较常见的高级语言有：**pascal** , **c 语言** , **basic** , fortran

把高级语言翻译成机器语言有两种方式：

编译方式：把高级语言写的源程序直接输入计算机，编译程序把整个源程序翻译成目标程序，然后由计算机执行目标程序（Pascal, C/C++ 采用编译方式）

解释方式：把高级语言写的源程序直接输入计算机，解释程序把源程序逐句翻译，翻译一句马上执行，边解释边执行

#### • 第四代：面向对象的高级语言

如果说第三代语言要求告诉计算机怎么做，那么第四代语言只需要告诉计算机做什么。

第一个面向对象的语言焦作 smalltalk。

比较常见的面向对象的高级语言有：free pascal , C/C++ , object pascal, c# 等

#### • 第五代：人工智能语言

代表语言是 LISP 语言和 Prolog 语言

#### • NOIP 比赛可使用的语言：**Pascal** , **C/C++**

### 5) 计算机的分类

#### 1. 以“代”分类：

- 第一代：1946~1957 年，**电子管计算机**
- 第二代：1958~1964 年，**晶体管计算机**
- 第三代：1965~1971 年，**中小规模集成电路计算机**
- 第四代：1972 至今，**大规模集成电路计算机**
- 第五代：**人工智能计算机**

## 2. 以功能规模分类

- 巨型机、大型机、中型机、小型机、微型机（如 PC 即个人电脑）

## 6) 计算机的应用

- 科学计算：弹道轨迹、天气预报、高能物理等等
- 数据处理（即信息处理）：企业管理、物资管理、电算化等
- 自动控制：工业自动化控制，卫星飞行方向控制
- 计算机辅助设计和教学：
  - 计算机辅助设计 CAD (Computer Aided Design)
  - 计算机辅助制造 CAM (Computer Aided Manufacturing)
  - 计算机辅助测试 CAT (Computer Aided Testing)
  - 计算机辅助教学 CAI (Computer Aided Instruction)
- 人工智能

## 第二部分 计算机硬件

1944 年，美籍匈牙利数学家 **冯·诺依曼** 提出计算机基本结构和工作方式的设想，为计算机的诞生和发展提供了理论基础。时至今日，尽管计算机软硬件技术飞速发展，但计算机本身的体系结构并没有明显的突破，当今的计算机仍属于 **冯·诺依曼架构**。

- 计算机硬件设备由存储器、运算器、控制器、输入设备和输出设备（I/O 设备）五个部分组成。
- 存储程序思想——把计算过程描述为由许多命令按一定顺序组成的程序，然后把程序和数据一起输入计算机，计算机对已存入的程序和数据处理后，输出结果。

### 1) 微型机的主要技术指标

- 字长：指计算机能够直接处理的二进制数据的位数。单位为位（BIT）
- 主频：指计算机主时钟在一秒钟内发出的脉冲数，在很大程度上决定了计算机的运算速度。
- 内存容量：是标志计算机处理信息能力强弱的一向技术指标。单位为字节（BYTE）。
- 外存容量：一般指软盘、硬盘、光盘。

计算机硬件由**五大部分**组成：**运算器、控制器、存储器、输入设备、输出设备（I/O 设备）**。



## 2) 中央处理器 (CPU——Central Processing Unit)

- 由**运算器、控制器**和一些**寄存器**组成;
- 运算器进行各种**算术运算和逻辑运算**;
- 控制器是计算机的指挥系统;
- CPU 的主要性能指标是主频和字长。

## 3) 存储器

### • 内部存储器

中央处理器能直接访问的存储器称为内部存储器，它包括快速缓冲存储器和主存储器，中央处理器不能直接访问的存储器称为外部存储器，外部存储器中的信息必须调入内存后才能为中央处理器处理。

主存储器：内存也常泛称主存，但严格上说，只有当内存中只有主存，而没有快速缓冲存储器时，才能称为主存。

主存储器按读写功能，可分**只读存储器 (ROM) [断电后信息不丢失]**和**随机存储器 (RAM) [断电后信息丢失]**两种。

### • 外部存储器

外存储器：也称为**辅助存储器**，一般容量较大，速度比主存较慢。

硬盘 (Hard disk)：将盘片、读写磁头及驱动装置精密地组装在一个密封盒里；采用接触式起停，非接触式读写的方式（磁盘不工作时，磁头停在磁盘表面的起停区，一旦加电后，磁头随着盘片旋转的气流“飞”起来，悬浮在磁盘表面，进行读写）。

软盘 (Floppy Disk)：目前常见的是 **3.5 英寸/1.44 MB** 的软盘。

光盘存储器 (CD-ROM)：普通的 CD-ROM，只能读，不能写；CD 盘片的存储量大约是 **650 MB**。

### • 高速缓冲存储器

又叫 **Cache**，由于 CPU 速度远大于内存，所以它在中间起协调作用。

### • 比较

内存和外存的主要区别：**内存小/外存大，内存贵/外存便宜，内存快/外存慢**  
存储器的速度：**寄存器>Cache>内存>外存（辅存）**

## 4) 输入设备

- 键盘 (Keyboard)：目前大多使用 104 或 108 键盘
- 鼠标 (Mouse)：主要有机械型鼠标和光电型鼠标两种
- 扫描仪 (Scanner)
- 手写笔、触摸屏、麦克风
- 视频输入设备、条形码扫描器

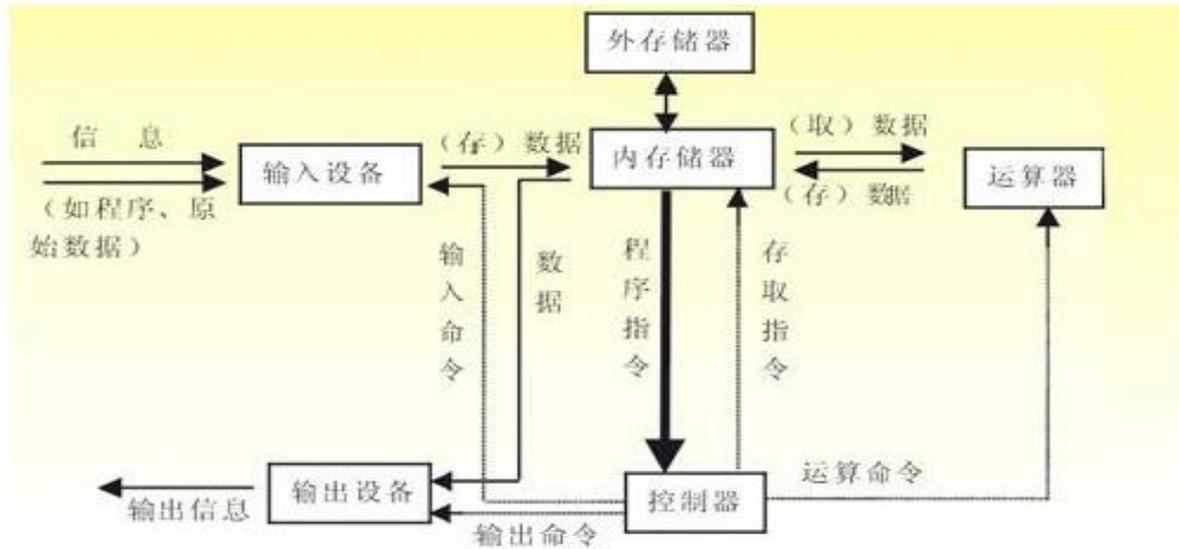
## 5) 输出设备

- 显示器 (Monitor)：目前主要有 CRT (阴极射线管) 显示器和 LCD 液晶显示器。
- 打印机 (Printer)：主要有针式打印机、喷墨打印机、激光打印机。
- 绘图仪、音箱

计算机在运行时，先从内存中取出一条指令，通过控制器的译码，按指令的要求，从存储器中取出数据进行指定的运算和逻辑操作等加工，然后再按地址把结果送到内存中去。

接下来取出下一条指令，在控制器的指挥下完成规定操作，依次进行下去，直到遇到停止指令

各部分联系示意图



### 第三部分 软件与操作系统

软件系统按其功能以及重要性可主要分为系统软件和应用软件两大类，系统软件管理整个计算机系统，应用软件在系统软件的基础上使计算机能解决各种实际问题。

#### 1) 应用软件在系统软件的基础上开发，但系统软件也属于应用软件

- 系统软件：用来支持应用软件的开发和运行的，主要是操作系统软件  
如：DOS、Windows95/98/2000、Unix、Linux、WindowsNT；
- 应用软件：为了某个应用目的而编写的软件，主要有文字处理软件、电子表格软件、数据库管理软件等。

#### 2) 操作系统 (OS—Operating System)

操作系统是控制与管理计算机系统资源的软件，是硬件的第一层扩充，任何应用软件的运行都必须依靠操作系统的支持。

操作系统大致有:dos , ucdos , windows 系列 , Linux , unix , os/2 ...

##### Windows 系列操作系统

Windows 是 Microsoft 公司开发的图形化界面的操作系统。

- 基本概念：  
图标、任务栏、标题栏、菜单栏、滚动条、工具栏、对话框、开始菜单……
- 基本操作：  
(1) 鼠标单击、双击、拖动，左键、右键功能；  
(2) 窗口操作：最大（小）化、大小调整、拖动、关闭、排列、切换；  
(3) 剪贴板：复制 (Ctrl+C)、粘贴 (Ctrl+V)、剪切 (Ctrl+X)  
复制屏幕图像：可将当前屏幕图形以 BMP 格式传送到剪贴板……  
(4) 其它：查找、运行、切换 Windows、进入 DOS 环境、文件夹选项

### 3) 语言处理软件

- Basic , Cobol , Pascal ,C/C++ ,Java

### 4) 数据库管理系统

- Sybase , **dbase** , foxbase , foxpro , **access** , **oracle** , **mysql**

### 5) 常用的应用软件

- Office 系列:Word(文字处理) , Excel (表格处理) ,Frontpage(网页制作) , Powerpoint(幻灯片)
- 图像处理软件: 画图 、 photoshop ,fireword , Acdsee
- 网页制作软件: dreamweaver
- 声音处理软件: goldwave
- 动画制作软件: flash

### 6) 常见的文件类型扩展名

- 可执行类文件: bat、 com、 exe、 sys、 tmp、 zip、 ……
- 文档类文件: doc、 xls、 txt、 htm、 ……
- 图像文件: bmp、 gif、 jpg、 psd、 ……
- 音频文件: wav、 avi、 mp3、 swf……

## 第四部分 进制与编码

### 1) 四种常用的数制及它们之间的相互转换:

进制	基数	基数个数	权	进数规律
十进制	0、1、2、3、4、5、6、7、8、9	10	$10^i$	逢十进一
二进制	0、1	2	$2^i$	逢二进一
八进制	0、1、2、3、4、5、6、7	8	$8^i$	逢八进一
十六进制	0、1、2、3、4、5、6、7、8、9、A、B、C、D、E、F	16	$16^i$	逢十六进一

十进制数转换为二进制数、八进制数、十六进制数的方法:

二进制数、八进制数、十六进制数转换为十进制数的方法: **按权展开求和法**

### 2) **二进制与十进制间的相互转换:**

#### (1) 二进制转十进制

方法: “**按权展开求和**”

$$\begin{aligned}\text{例: } (1011.01)_2 &= (1 \times 2^3 + 0 \times 2^2 + 1 \times 2^1 + 1 \times 2^0 + 0 \times 2^{-1} + 1 \times 2^{-2})_{10} \\ &= (8 + 0 + 2 + 1 + 0 + 0.25)_{10} \\ &= (11.25)_{10}\end{aligned}$$

规律: 个位上的数字的次数是0, 十位上的数字的次数是1, …… , 依次递增, 而十分位的数字的次数是-1, 百分位上数字的次数是-2, …… , 依次递减。

**注意: 不是任何一个十进制小数都能转换成有限位的二进制数。**

## (2) 十进制转二进制

十进制整数转二进制数：“除以 2 取余，逆序排列”（短除反取余法）

例：  $(89)_{10} = (1011001)_2$

2	89	
2	44	.....1
2	22	.....0
2	11	.....0
2	5	.....1
2	2	.....1
2	1	.....0
	0	.....1

十进制小数转二进制数：“乘以 2 取整，顺序排列”（乘 2 取整法）

例：  $(0.625)_{10} = (0.101)_2$

	0.625	
X	2	
	1.25	1
X	2	
	0.5	0
X	2	
	1.0	1

## 3) 八进制与二进制的转换：

(1) 二进制数转换成八进制数：从小数点开始，整数部分向左、小数部分向右，每 3 位为一组用一位八进制数的数字表示，不足 3 位的要用“0”补足 3 位，就得到一个八进制数。

(2) 八进制数转换成二进制数：把每一个八进制数转换成 3 位的二进制数，就得到一个二进制数。

例：将八进制的 37.416 转换成二进制数：

3   7   .   4   1   6  
011 111 . 100 001 110

即：  $(37.416)_8 = (11111.100001110)_2$

例：将二进制的 10110.0011 转换成八进制：

0 1 0 1 1 0 . 0 0 1 1 0 0  
2        6        .        1        4

即：  $(10110.011)_2 = (26.14)_8$

## 4) 十六进制与二进制的转换：

(1) 二进制数转换成十六进制数：从小数点开始，整数部分向左、小数部分向右，每 4 位为一组用一位十六进制数的数字表示，不足 4 位的要用“0”补足 4 位，就得到一个十六进制数。

(2) 十六进制数转换成二进制数：把每一个八进制数转换成 4 位的二进制数，就得到一个二进制数。

例：将十六进制数 5DF.9 转换成二进制：

5     D     F     .     9  
0101 1101 1111 . 1001

即：  $(5DF.9)_{16} = (10111011111.1001)_2$

例：将二进制数 110001.111 转换成十六进制：

$$\begin{array}{ccc} 0110 & 0001 & . & 1110 \\ 6 & 1 & . & E \end{array}$$

即：(110001.111)<sub>2</sub> = (61.E)<sub>16</sub>

注意：以上所说的二进制数均是无符号的数。这些数的范围如下表：

无符号位二进制数位数	数值范围	十六进制范围表示法
8 位二进制数	0~255 (255=2 <sup>8</sup> -1)	00~0FFH
16 位二进制数	0~65535 (65535=2 <sup>16</sup> -1)	0000H~0FFFFH
32 位二进制数	0~2 <sup>32</sup> -1	00000000H~FFFFFFFFH

### 5) 带符号数的机器码表示方法

#### (一) 带符号二进制数的表示方法：

带符号二进制数用最高位的一位数来表示符号：0 表示正，1 表示负。

含符号位二进制数位数	数值范围	十六进制范围表示法
8 位二进制数	-128 ~ +127	80H~7FH
16 位二进制数	-32768 ~ +32767	8000H~7FFFH
32 位二进制数	-2147483648 ~ +2147483647	80000000H~7FFFFFFFH

#### (二) 符号位的表示：最常用的表示方法有原码、反码和补码。

**(1) 原码表示法：**一个机器数 x 由符号位和有效数值两部分组成，设符号位为 x<sub>0</sub>，x 真值的绝对值 |x|=x<sub>1</sub>x<sub>2</sub>x<sub>3</sub>...x<sub>n</sub>，则 x 的机器数原码可表示为：

$$[x]_{原} = x_0x_1x_2...x_n, \text{ 当 } x \geq 0 \text{ 时, } x_0=0, \text{ 当 } x < 0 \text{ 时, } x_0=1.$$

例如：已知：x<sub>1</sub>=-1011B, x<sub>2</sub>=+1001B, 则 x<sub>1</sub>, x<sub>2</sub> 有原码分别是

$$[x_1]_{原}=11011B, [x_2]_{原}=01001B$$

规律：正数的原码是它本身，负数的原码是取绝对值后，在最高位（左端）补“1”。

**(2) 反码表示法：**一个负数的原码符号位不变，其余各位按位取反就是机器数的反码表示法。正数的反码与原码相同。

按位取反的意思是该位上是 1 的，就变成 0，该位上是 0 的就变成 1。即  $\bar{1}=0, \bar{0}=1$

例：x<sub>1</sub> = -1011B, x<sub>2</sub> = +1001B, 求 [x<sub>1</sub>]<sub>反</sub> 和 [x<sub>2</sub>]<sub>反</sub>。

解：[x<sub>1</sub>]<sub>反</sub> = 10100B, [x<sub>2</sub>]<sub>反</sub> = 01001B

#### **(3) 补码表示法：**首先分析两个十进制数的运算：79-38=41, 79+62=141

如果使用两位数的运算器，做 79+62 时，多余的 100 因为超出了运算器两位数的范围而自动丢弃，这样在做 79-38 的减法时，用 79+62 的加法同样可以得到正确结果。

模是批一个计量系统的测量范围，其大小以计量进位制的基数为底数，位数为指数的幂。如两位十进制数的测量范围是 1—9，溢出量是 100，模就是 10<sup>2</sup>=100，上述运算称为模运算，可以写作：

$$79 + (-38) = 79 + 62 \pmod{100}$$

进一步写为 -38=62，此时就说 -38 的补法（对模 100 而言）是 62。计算机是一种有限字长的数字系统，因此它的运算都是有模运算，超出模的运算结果都将溢出。n 位二进制的模是 2<sup>n</sup>，

一个数的补码记作 $[x]_{\text{补}}$ ，设模是 $M$ ， $x$ 是真值，则补码的定义如下：

$$[x]_{\text{补}} = \begin{cases} [x]_{\text{原}} & (x \geq 0) \\ M + x & (x < 0) \end{cases}$$

例：设字长 $n=8$ 位， $x=-1011011\text{B}$ ，求 $[x]_{\text{补}}$ 。

解：因为 $n=8$ ，所以模 $M=2^8=10000000\text{B}$ ， $x<0$ ，所以

$$[x]_{\text{补}} = M + x = 10000000\text{B} - 1011011\text{B} = 10100101\text{B}$$

注意：这个 $x$ 的补码的最高位是“1”，表明它是一个负数。对于二进制数还有一种更加简单的方法由原码求出补码：

#### (4) 总结

- a) 正数的原码、反码、补码都是相同的；
- b) 负数的补码是将原码符号位保持“1”之后，其余各位按位取反，末位再加1便得到补码，即取其原码的反码再加“1”： $[x]_{\text{补}} = [x]_{\text{反}} + 1$ 。

下表列出 $\pm 0$ ， $\pm 39$ ， $\pm 127$ 以及 $-128$ 的8位二进制原码，反码和补码并将补码用十六进制表示。

真值	原码 (B)	反码 (B)	补码 (B)	补码 (H)
+127	0 111 1111	0 111 1111	0 111 1111	7F
+39	0 010 0111	0 010 0111	0 010 0111	27
+0	0 000 0000	0 000 0000	0 000 0000	00
-0	1 000 0000	1 111 1111	0 000 0000	00
-39	1 010 0111	1 101 1000	1 101 1001	D9
-127	1 111 1111	1 000 0000	1 000 0001	81
-128	无法表示	无法表示	1 000 0000	80

从上可看出，真值+0和-0的补码表示是一致的，但在原码和反码表示中具有不同形式。8位补码机器数可以表示-128，但不存在+128的补码与之对应，由此可知，8位二进制补码能表示数的范围是-128——+127。还要注意，不存在-128的8位原码和反码形式。

## 6) 定点数和浮点数

### (1) 定点数 (Fixed-Point Number)

计算机处理的数据不仅有符号，而且大量的数据带有小数，小数点不占有二进制一位而是隐含在机器数里某个固定位置上。通常采取两种简单的约定：一种是约定所有机器数的小数的小数点位置隐含在机器数的最低位之后，叫定点纯整机器数，简称定点整数。另一种约定所有机器数的小数点隐含在符号位之后、有效部分最高位之前，叫定点纯小机器数，简称定点小数。无论是定点整数，还是定点小数，都可以有原码、反码和补码三种形式。

### (2) 浮点数 (Floating-Point Number)

计算机多数情况下采作浮点数表示数值，它与科学计数法相似，把一个二进制数通过移动小数点位置表示成阶码和尾数两部分：

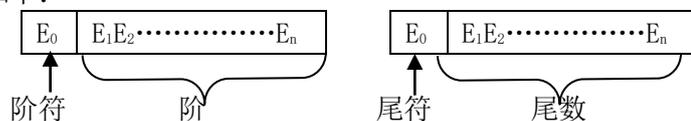
$$N = 2^E \times S$$

其中： $E$ —— $N$ 的阶码 (Exponent)，是有符号的整数

$S$ —— $N$ 的尾数 (Mantissa)，是数值的有效数字部分，一般规定取二进制定点纯小数形式。

例： $1011101\text{B} = 2^7 * 0.1011101$ ， $101.1101\text{B} = 2^{+3} * 0.1011101$ ， $0.01011101\text{B} = 2^{-1} * 0.1011101$

浮点数的格式如下：



浮点数由阶码和尾数两部分组成，底数 2 不出现，是隐含的。阶码的正负符号  $E_0$ ，在最前位，阶反映了数  $N$  小数点的位置，常用补码表示。二进制数  $N$  小数点每左移一位，阶增加 1。尾数是这点小数，常取补码或原码，码制不一定与阶码相同，数  $N$  的小数点右移一位，在浮点数中表现为尾数左移一位。尾数的长度决定了数  $N$  的精度。尾数符号叫尾符，是数  $N$  的符号，也占一位。

例：写出二进制数  $-101.1101B$  的浮点数形式，设阶码取 4 位补码，尾数是 8 位原码。

$$-101.1101 = -0.1011101 \times 2^3$$

浮点形式为：

阶码 0011 尾数 11011101

补充解释：阶码 0011 中的最高位“0”表示指数的符号是正号，后面的“011”表示指数是“3”；尾数 11011101 的最高位“1”表明整个小数是负数，余下的 1011101 是真正的尾数。

例：计算机浮点数格式如下，写出  $x=0.0001101B$  的规格化形式，阶码是补码，尾数是原码。

$$x=0.0001101=0.1101 \times 10^{-3}$$

$$\text{又 } [-3]_{\text{补}} = [-001B]_{\text{补}} = [1011]_{\text{补}} = 1101B$$

所以浮点数形式是

1	101	0	1101000
---	-----	---	---------

## 7) ASCII 码 (American Standard Code for Information Interchange)

### 美国标准信息交换标准码

常用字符有 128 个，编码从 0 到 127.

空格	—	—	32
‘0’ ~ ‘9’	—	—	48 ~ 57
‘A’ ~ ‘Z’	—	—	65 ~ 90
‘a’ ~ ‘z’	—	—	97 ~ 122

**ASCII 码用一个字节来表示，所以英文字符大小都是一个字节**

**汉字在计算机内使用机内码存储，需要两个字节**

## 8) 汉字信息编码

### 1. 汉字输入码

汉字输入方法大体可分为：区位码（数字码）、音码、形码、音形码。

(1) 区位码：优点是无重码或重码率低，缺点是难于记忆；

(2) 音码：优点是大多数人都易于掌握，但同音字多，重码率高，影响输入的速度；

(3) 形码：根据汉字的字型进行编码，编码的规则较多，难于记忆，必须经过训练才能较好地掌握；重码率低；

(4) 音形码：将音码和形码结合起来，输入汉字，减少重码率，提高汉字输入速度。

### 2. 汉字交换码

汉字交换码是指不同的具有汉字处理功能的计算机系统之间在交换汉字信息时所使用的代码标准。自国家标准 GB2312-80 公布以来，我国一直沿用该标准所规定的国标码作为统一的汉字信息交换码。

GB2312-80 标准包括了 **6763 个汉字**，按其使用频度分为一级汉字 **3755 个**和二级汉字 **3008 个**。一级汉字按拼音排序，二级汉字按部首排序。此外，该标准还包括标

点符号、数种西文字母、图形、数码等符号 682 个。

由于 GB2312-80 是 80 年代制定的标准，在实际应用时常常感到不够，所以，建议处理文字信息的产品采用新颁布的 GB18030 信息交换用汉字编码字符集，这个标准繁、简字均处同一平台，可解决两岸三地间 GB 码与 BIG5 码间的字码转换不便的问题。

### 3. 字形存储码

字形存储码是指供计算机输出汉字（显示或打印）用的二进制信息，也称字模。通常，采用的是数字化点阵字模。

一般的点阵规模有  $16 \times 16$ ， $24 \times 24$ ， $32 \times 32$ ， $64 \times 64$  等，每一个点在存储器中用一个二进制位 (bit) 存储。例如，在  $16 \times 16$  的点阵中，需  $16 \times 16 \text{bit} = 32 \text{ byte}$  的存储空间。在相同点阵中，不管其笔划繁简，每个汉字所占的字节数相等。

为了节省存储空间，普遍采用了字形数据压缩技术。所谓的矢量汉字是指用矢量方法将汉字点阵字模进行压缩后得到的汉字字形的数字化信息。

- 图像的大小表示为：分辨率\*像素大小

例如  $1024 * 768$  分辨率下，一张 16 位的图片大小 =  $1024 * 768 * 16 / 8 (B)$

## 9) 常见的单位与换算：

二进制位：简称位，计算机中信息表示的最小单位，用 bit（比特）表示，也就是一个 0 或者一个 1。例如 8 可以表示为四位的二进制 1000

**8BIT = 1BYTE (简称 B)**

**1024B = 1KB (千字节)**

**1024KB = 1MB (兆字节)**

**1024MB = 1GB (千兆字节)**

**1024GB = 1TB (兆兆字节)**

## 10) 逻辑运算

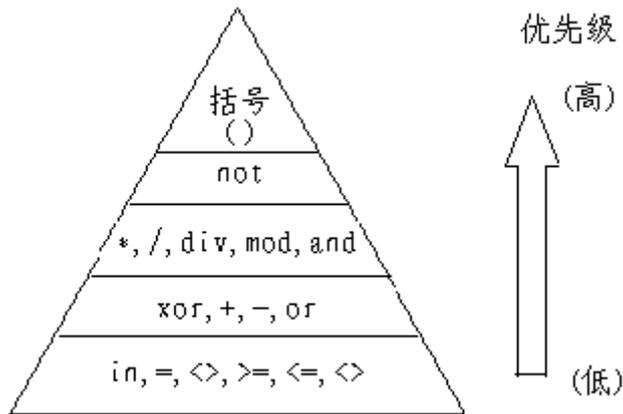
概念介绍：

非：not  $\neg$     与：and  $\wedge$     或：or  $\vee$     异或：xor  $\oplus$

运算级比较：

括号 > 非 > 与 > 或、异或    ( or 和 xor 是同级的)

如果加入加减乘除，就是以下这样：



注意：同级的运算符不分高低，计算时按照从左到右运算。  
运算法则：

- ∧：两边相同返回真；两边中有一个不同则返回假；  
A and B 只要 A 和 B 有一个是 false 就是 false
- ∨：两边只要又一边为真，即可返回真；否则返回假；  
A or B 只要 A 和 B 有一个是 true 就是 true
- ⊕：两边一样则假，不一样则真  
A xor B 只要 A 和 B 不一样就是 true，相同就是 true  
即 A xor A = 0
- ¬：如果后面的为假，则返回真，如果为真，则返回假；  
Not 表示取反，Not true = false

A	B	Not A	A and B	A or B	A xor B
False	False	True	False	False	False
False	True	True	False	True	True
True	False	False	False	True	True
True	True	False	True	True	False

## 11) 位运算

表示把整数化成二进制后逐位比较，求出比较结果后，再转换成十进制  
常见运算符：and, or, xor, >>(右移), <<(左移)

运算	10 and 6	10 or 6	10 xor 6	10 >> 3	10 << 3
10	1010	1010	1010	1010 右移 三位	1010 左移 三位
6	0110	0110	0110		
二进制结果	0010	1110	1100	1	1010000
最后结果	2	14	12	1	80
说明	转换成二进制后，位数不够的在高位补 0， 然后逐位比较 1 相当于 true 0 相当于 false			$10 \gg 3 = 10 / 8 = 1$ $10 \ll 3 = 10 * 8 = 80$ $N \gg X = N / (2^X)$ $N \ll X = N * (2^X)$	

## 12) 表达式的三种类型

(一) 什么是中缀、前缀、后缀表达式

例子：有一个数学表达式： $a*(b+c)-d$

1: **中缀表达式**

$a*(b+c)-d$  这个是中缀表达式，运算符在中间，我们数学中的表达式都是中缀表达式

2: **后缀表达式**

计算机运算时采用的是后缀表达式，这个表达式的后缀表达式是  $abc+*d-$   
计算机从左到右进行运算，遇到运算符就把前面两个数值进行相关计算。

3: **前缀表达式**

前缀表达式就是运算符在前面，数值放在后面，计算机计算时从右往左运算。  
遇到运算符就把前面两个数值进行相关计算。这个表达式的前缀表达式是  $-*a+bcd$

(二) 把中缀表达式转化成后缀表达式的思想如下：

算法基本思想：

使用三个数组，一个数组保存用户输入的表达式（中缀表达式），一个数组保存后缀表达式，一个数组作为运算符的栈。

从头到尾扫描中缀表达式，对不同类型的字符按不同情况处理：

- 1、如果是数字则直接放入后缀表达式数组；
- 2、如果是左括号则直接入栈；
- 3、如果是右括号，则把从栈顶直到对应左括号之间的运算符依次退栈，写入后缀表达式数组，并清除对应的左括号；
- 4、对于运算符，如果该运算符的优先级大于栈顶优先级，则直接入栈，若该运算符的优先级小于等于栈顶优先级，则把栈顶运算符逐个出栈，写入后缀表达式数组，直到遇到左括号或者遇到比它优先级小的运算符。然后再入栈；
- 5、扫描完成后，取出栈中所有运算符，写入后缀表达式数组。

## 第五部分 信息安全

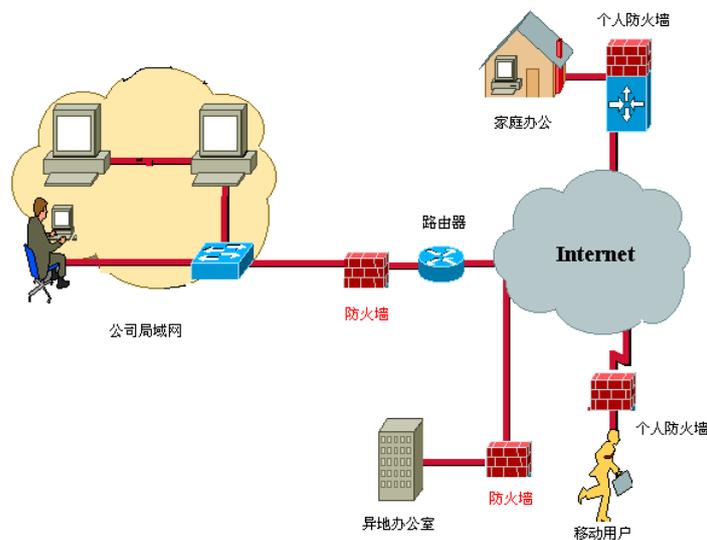
计算机安全 (computer security) 是指防范与保护计算机系统及其信息资源在生存过程中免受蓄意攻击、人为失误和自然灾害等引起的损失和破坏。

计算机病毒是人类自己想像和发明出来的，它是一种特殊的程序，有着与生物病毒极为相似的特点。一是寄生性，它们大多依附在别的程序上面。二是隐蔽性，它们是悄然进入系统的，人们很难察觉。三是潜伏性，它们通常是潜伏在计算机程序中，只在一定条件下才发作的。四是传染性，它们能够自我复制繁殖，通过传输媒介蔓延。五是破坏性，轻则占用一定数量的系统资源，重则破坏整个系统。

对于计算机病毒，我们不必谈虎变色，而应采取积极的防治态度。首先，要防止“病从口入”，因为病毒不是自生的，而是外来的。另外，要用优秀的防杀病毒软件，对外来

的软件和资料要进行严格的检查和杀毒。注意，防杀病毒软件需要及时更新(主要是其中的数据文件)，一般每周一次，不更新基本上等于没有防杀毒功能。

20世纪50、60年代，黑客(hacker)曾是编程高手的代名词。后来，黑客成为一个独特的群体，他们通过各种渠道交流技艺，不少人以攻击计算机及其网络系统为乐趣。黑客们的胆大妄为已经给社会造成了很大的影响，一些黑客已经蜕变为威胁社会安全的罪犯。要防止“黑客”攻击，主要方法是加强安全措施，例如设置防火墙(见图3.1.1)。防火墙是一种计算机设备，它设置在内部网络与外部网络之间，起一个隔离的作用，既可以阻止外部信息非法进入内部系统，也可以阻止内部人员非法访问外部系统。



## ● 例题

1. 计算机病毒传染的必要条件是( B )。

- A) 在内存中运行病毒程序                      **B) 对磁盘进行读写操作**  
C) 在内存中运行含有病毒的程序              D) 复制文件

2. 计算机病毒是( B )

- A) 通过计算机传播的危害人体健康的一种病毒  
**B) 人为制造的能够侵入计算机系统并给计算机带来故障的程序或指令集合**  
C) 一种由于计算机元器件老化而产生的对生态环境有害的物质  
D) 利用计算机的海量高速运算能力而研制出来的用于疾病预防的新型病毒

3、计算机病毒的特点是( C )

- A. 传播性、潜伏性、易读性与隐蔽性              B. 破坏性、传播性、潜伏性与安全性  
**C. 传播性、潜伏性、破坏性与隐蔽性**              D. 传播性、潜伏性、破坏性与易读性

## 第六部分 计算机网络

### 1) 关于网络的一些定义

所谓计算机网络，就是利用通信线路和设备，把分布在不同地理位置上的多台计算机连接起来。

计算机网络是现代通信技术与计算机技术相结合的产物。网络中计算机与计算机之间的通信依靠协议进行。协议是计算机收、发数据的规则。

- **TCP/IP**

用于网络的一组通讯协议。包括 IP(Internet Protocol)和 TCP(Transmission Control Protocol)。

TCP/IP是一组协议，包括上百个各种功能的协议，其中TCP 和IP是最核心的两个协议。TCP/IP 协议把Internet网络系统描述成具有四个层次功能的网络模型。

1. **链路层**：这是TCP/IP 结构的第一层，也叫网络接口层，其功能是提供网络相邻节点间的信息传输以及网络硬件和设备驱动。

2. **网络层**：（IP协议层）其功能是提供源节点和目的节点之间的信息传输服务，包括寻址和路由器选择等功能。

3. **传输层**：（TCP 协议）其功能是提供网络上的各应用程序之间的通信服务。

4. **应用层**：这是TCP/IP最高层，其功能是为用户提供访问网络环境的手段，主要提供FTP、TELNET、GOPHER等功能软件。

IP协议适用于所有类型网络。TCP 协议则处理IP协议所遗留的通信问题，为应用程序提供可靠的通信连接，并能自动适应网络的变化。TCP/IP 目前成为最为成功的网络体系结构和协议规范。

- Netbeui

一种非常简单的协议，MICROSOFT 开发。

- IPX

用于 NOVELL 网络。

- 网络的发展

计算机网络的发展过程大致可以分为三个阶段：

{	远程终端联机阶段：	主机—终端
	计算机网络阶段：	计算机—计算机
	Internet 阶段：	Internet

### 2) 网络的主要功能：

- (1) 资源共享
- (2) 信息传输
- (3) 分布处理
- (4) 综合信息服务

### 3) 网络的分类

计算机网络的分类方式有很多种，可以按**地理范围**、**拓扑结构**、**传输速率**和**传输介质**等分类。

#### 1. 按地理范围分类

##### ① **局域网LAN**(Local Area Network)

局域网地理范围一般几百米到10km 之内，属于小范围内的连网。如一个建筑物内、一个学校内、一个工厂的厂区内等。局域网的组建简单、灵活，使用方便。

② **城域网MAN**(Metropolitan Area Network)

城域网地理范围可从几十公里到上百公里,可覆盖一个城市或地区,是一种中等形式的网络。

③ **广域网WAN**(Wide Area Network)

广域网地理范围一般在几千公里左右,属于大范围连网。如几个城市,一个或几个国家,是网络系统中的最大型的网络,能实现大范围的资源共享,如国际性的Internet 网络。

## 2. 按传输速率分类

① 网络的传输速率有快有慢,传输速率快的称**高速网**,传输速率慢的称**低速网**。

传输速率的单位是b/s(每秒比特数,英文缩写为bps)。

一般将传输速率在Kb/s—Mb/s范围的网络称低速网,在Mb/s—Gb/s 范围的网称高速网。

也可以将Kb/s 网称低速网,将Mb/s网称中速网,将Gb/s网称高速网。

② 网络的传输速率与网络的带宽有直接关系。带宽是指传输信道的宽度,带宽的单位是Hz(赫兹)。按照传输信道的宽度可分为**窄带网**和**宽带网**。

一般将KHz—MHz带宽的网称为窄带网,将MHz—GHz 的网称为宽带网,也可以将kHz 带宽的网称窄带网,将MHz 带宽的网称中带网,将GHz 带宽的网称宽带网

③ **通常情况下,高速网就是宽带网,低速网就是窄带网。**

## 3. 按传输介质分类

传输介质是指数据传输系统中发送装置和接受装置间的物理媒体,按其物理形态可以划分为有线和无线两大类。

① 有线网

传输介质采用有线介质连接的网络称为有线网,常用的有线传输介质有**双绞线**、**同轴电缆**和**光导纤维**。

- 双绞线是由两根绝缘金属线互相缠绕而成,这样的一对线作为一条通信线路,由四对双绞线构成双绞线电缆。双绞线点到点的通信距离一般不能超过100m。目前,计算机网络上使用的双绞线按其传输速率分为三类线、五类线、六类线、七类线,传输速率在10Mbps到600Mbps之间,双绞线电缆的连接器一般为RJ-45。
- 同轴电缆由内、外两个导体组成,内导体可以由单股或多股线组成,外导体一般由金属编织网组成。内、外导体之间有绝缘材料,其阻抗为50Ω。同轴电缆分为粗缆和细缆,粗缆用DB-15连接器,细缆用BNC和T 连接器。
- 光缆由两层折射率不同的材料组成。内层是具有高折射率的玻璃单根纤维体组成,外层包一层折射率较低的材料。光缆的传输形式分为单模传输和多模传输,单模传输性能优于多模传输。所以,光缆分为单模光缆和多模光缆,单模光缆传送距离为几十公里,多模光缆为几公里。光缆的传输速率可达到每秒几百兆位。光缆用ST 或SC 连接器。光缆的优点是不会受到电磁的干扰,传输的距离也比电缆远,传输速率高。光缆的安装和维护比较困难,需要专用的设备。

② 无线网

采用无线介质连接的网络称为无线网。目前无线网主要采用三种技术:微波通信,红外线通信和激光通信。这三种技术都是以大气为介质的。其中微波通信用途最广,目前的卫星网就是一种特殊形式的微波通信,它利用地球同步卫星作中继站来转发微波信号,一个同步卫星可以覆盖地球的三分之一以上表

面,三个同步卫星就可以覆盖地球上全部通信区域。

#### 4. 按拓扑结构分类

计算机网络的物理连接形式叫做网络的物理拓扑结构。连接在网络上的计算机、大容量的外存、高速打印机等设备均可看作是网络上的一个节点,也称为工作站。计算机网络中常用的拓扑结构有总线型、星型、环型等。

##### ① 总线拓扑结构

总线拓扑结构是一种共享通路的物理结构。这种结构中总线具有信息的双向传输功能,普遍用于局域网的连接,总线一般采用同轴电缆或双绞线。

**总线拓扑结构的优点是:** 安装容易,扩充或删除一个节点很容易,不需停止网络的正常工作,节点的故障不会殃及系统。由于各个节点共用一个总线作为数据通路,信道的利用率高。但总线结构也有其缺点:由于信道共享,连接的节点不宜过多,并且总线自身的故障可以导致系统的崩溃。

##### ② 星型拓扑结构

星型拓扑结构是一种以中央节点为中心,把若干外围节点连接起来的辐射式互联结构。这种结构适用于局域网,特别是近年来连接的局域网大都采用这种连接方式。这种连接方式以双绞线或同轴电缆作连接线路。

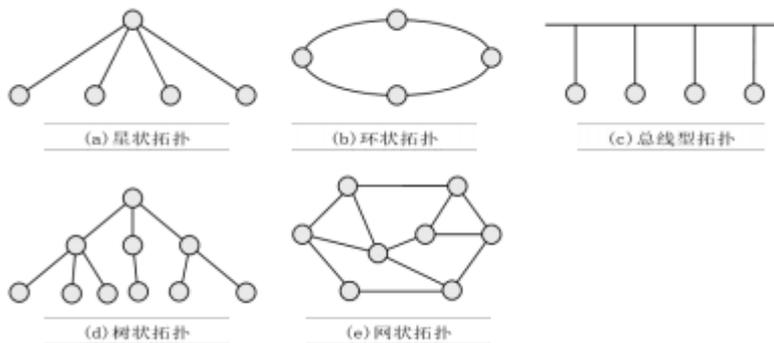
**星型拓扑结构的特点是:** 安装容易,结构简单,费用低,通常以集线器(Hub)作为中央节点,便于维护和管理。中央节点的正常运行对网络系统来说是至关重要的。

##### ③ 环型拓扑结构

环型拓扑结构是将网络节点连接成闭合结构。信号顺着从一个方向从一台设备传到另一台设备,每一台设备都配有一个收发器,信息在每台设备上的延时间是固定的。

这种结构特别适用于实时控制的局域网系统。

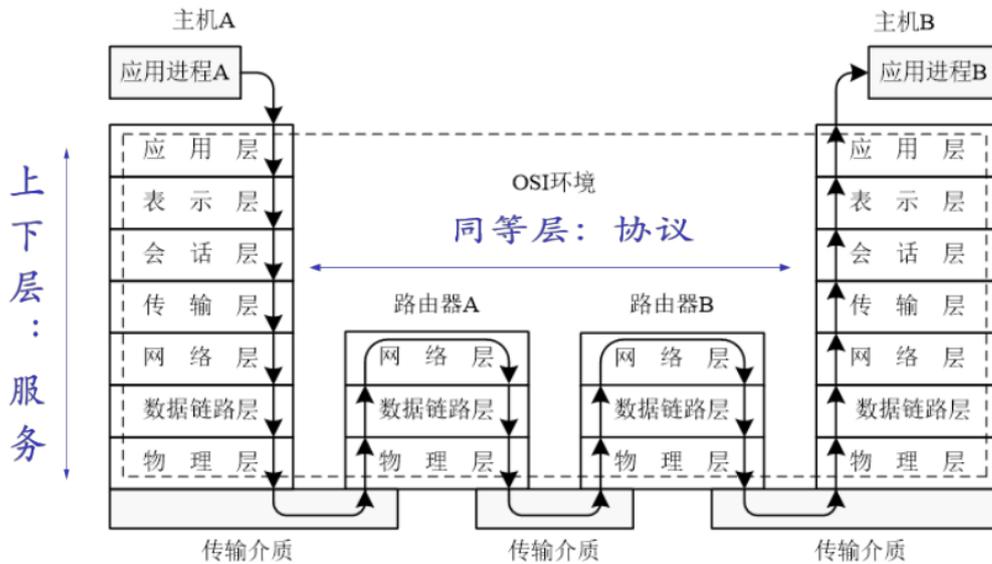
**环型拓扑结构的特点是:** 安装容易,费用较低,电缆故障容易查找和排除。有些网络系统为了提高通信效率和可靠性,采用了双环结构,即在原有的单环上再套一个环,使每个节点都具有两个接收通道。环型网络的弱点是,当节点发生故障时,整个网络就不能正常工作。



#### 4) 网络的体系结构

OSI 的七层体系结构从上往下依次是:

应用层,表示层,会话层,运输层,网络层,数据链路层,物理层



## 5) 局域网的工作方式

1. 客户机/服务器 (Client/Server):  
提供资源并管理资源的计算机称为服务器；使用共享资源的计算机称客户机；
2. 对等 (Peer-to-Peer):  
不使用服务器来管理网络共享资源，所以的计算机处于平等的地位。

## 6) Internet 的形成与发展

Internet 又称国际互联网，规范的译名是“因特网”，指当前各国、各地区众多开发的网络连接在一起而形成的全球性网络。

我国 Internet 的发展情况：八十年代末，九十年代初才起步。

1989 年我国第一个公用分组交换网 CNPAC 建成运行。

我国已陆续建成与 Internet 互联的四个全国范围的公用网络：

中国公用计算机互联网 (CHINANET)、中国金桥信息网 (CHINAGBN)  
中国教育和科研计算机网 (CERNET)、中国科学技术网 (CSTNET)

## 7) IP 地址：

我们把整个 Internet 看作一个单一的、抽象的网络，所谓 IP 地址，就是为 Internet 中的每一台主机分配一个在全球范围唯一地址。**IPv4 地址是由 32 位二进制数码表示的，为方便记忆，把这 32 位二进制数每 8 个一段用“.” 隔开，再把每一段的二进制数化成十进制数，也就得到我们现在所看到的 IP 地址形式。**

**IP 地址是用“.” 隔开地四个十进制整数，每个数字取值为 0—255。**

IP 地址分 A、B、C、D；E 五类，目前大量使用的是 A、B、C 三类，D 类为 Internet 体系结构委员会 IAB 专用，E 类保留在今后使用。

**最高位 1.. 126 为 A 类，128.. 191 是 B 类，192.. 223 是 C 类。**

- **域名**

域名地址采用层次结构，一个域名一般有 3—5 个子段，中间用“.” 隔开。IP 地址作为 Internet 上主机的数字标识，对计算机网络来说是非常有效的。但对于使用者来说，很难记忆这些由数字组成的 IP 地址了。为此，人们研究出一种字符型标识，在

Internet上采用“名称”寻址方案，为每台计算机主机都分配一个独有的“标准名称”，这个用字符表示的“标准名称”就是我们现在所广泛使用的域名（DN，domain name）。因此主机的域名和IP地址一样，也采用分段表示的方法。其结构一般是如下样式：计算机名.组织结构名.网络名.最高层域名。

顶级域名有**三类**：

- **国家顶级域名**，如 cn（中国）、us（美国）、uk（英国）；
- **国际顶级域名**—— int，国际性组织可在 int 下注册；
- **通用顶级域名**，如：com、net、edu、gov、org、……

有了域名标识，对于计算机用户来说，在使用上的确方便了很多。但计算机本身并不能自动识别这些域名标识，于是域名管理服务器DNS（domain name system）就应运而生了。所谓的域名管理系统DNS（domain name system）就是以主机的域名来代替其在Internet上实际的IP地址的系统，它负责将Internet上主机的域名转化为计算机能识别的IP地址。从DNS的组织结构来看，它是一个按照层次组织的分布式服务系统；从它的运行机制来看，DNS更像一个庞大的数据库，只不过这个数据库并不存储在任一计算机上，而是分散在遍布于整个Internet上数以千计的域名服务器中而已。通过上面的IP地址、域名DN和域名管理系统DNS，就把Internet上面的每一台主机给予了唯一的定位。三者之间的具体联系过程如下：当连接网络并输入想访问主机的域名后，由本地机向域名服务器发出查询指令，域名服务器通过连接在整个域名管理系统查询对应的IP地址，如找到则返回相应的IP地址，反之则返回错误信息。说到这里，想必大家都明白了为什么当我们在浏览时，浏览器左下角的状态条上会有这样的信息：“正在查找xxxxxx”、“xxxxxx已经发现，正在连接xxxxxx”，其实这也就是域名通过DNS转化为IP地址的过程。

当然域名通过DNS转化为IP地址需要等待一段时间，因为如果你所使用的域名服务器上如果没有你所需要域名的对应IP地址，它就会向上级域名服务器查询，如此类推，直至查到结果，或返回无效信息。一般而言，这个查询过程都非常短，你很难察觉到。

## 8) **Internet（译为因特网或国际互联网）的服务与工具**

Internet的服务有：电子邮件、远程登陆、文件传输、信息服务等；

**电子邮件（E-Mail）**：电子邮件地址格式为：

收信人邮箱名@邮箱所在主机的域名。例：winner01@21cn.com，  
qfit168@yahoo.com.cn

**远程登陆（Telnet）**：指通过Internet与其它主机连接。

登陆上另一主机，你就可以使用该主机对外开放的各种资源，如联机检索、数据查询。

- **文件传输（FTP）**：用于在计算机间传输文件。如下载软件等。

## 9) **全球信息网（www—World Wide Web）**

又称**万维网**，是一个全球规模的信息服务系统，由遍布于全世界的数以万计的Web站点组成。