

计算机体系结构

第二讲

计算机科学与技术学院

舒燕君

Recap

- 课程基本介绍
- 课程的学习目的
- 课程的学习内容
 - ✓ 现代计算机体系结构的基本概念
 - ✓ 现代计算机体系结构的设计思想和评价技术
 - ✓ 典型计算机系统的基本结构及其工作原理等
- 计算机体系结构概念的演变

第 0 章 前言

第 1 章 计算机体系结构的基本概念

第 2 章 指令系统

第 3 章 流水线技术

第 4 章 指令级并行

第 5 章 存储层次

第 6 章 输入输出系统

程序员所看到的计算机的属性

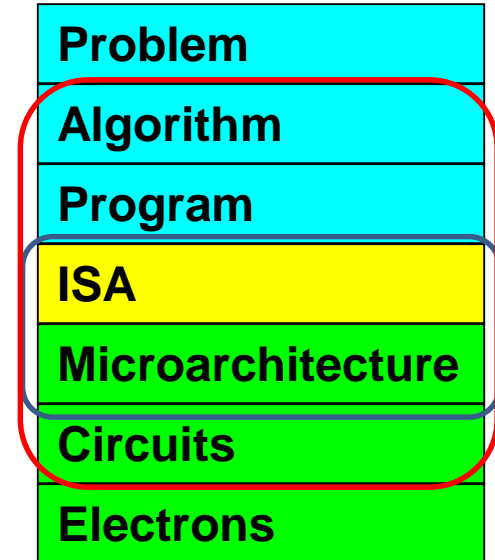
- 对于通用寄存器型机器，这些属性主要是指：
 - (1) 数据表示：硬件能直接辨认和处理的数据类型
 - (2) 寻址规则：最小寻址单元、寻址方式及其表示
 - (3) 寄存器定义：寄存器的定义、数量和使用方式
 - (4) 指令系统：机器指令的操作类型和格式、指令间的排序和控制机构等
 - (5) 中断系统：中断的类型和中断响应硬件的功能等
 - (6) 机器工作状态的定义和切换：如管态和目态等
 - (7) 存储系统：程序员可用的最大存储容量
 - (8) 信息保护：信息保护方式和硬件的支持
 - (9) I/O结构：I/O寻址方式、数据传送的方式等

What is Computer Architecture?

- **ISA+Implementation definition:** The science and art of designing, selecting, and interconnecting hardware components and designing the hardware/software interface to create a computing system that meets functional, performance, energy consumption, cost, and other specific goals.
- **Traditional (ISA-only) definition:** “The term *architecture* is used here to describe the attributes of a system as seen by the programmer, i.e., the conceptual structure and functional behavior as distinct from the organization of the dataflow and controls, the logic design, and the physical implementation.” *Gene Amdahl*, IBM Journal of R&D, April 1964

ISA vs. Microarchitecture

- ISA
 - Agreed upon interface between software and hardware
 - SW/compiler assumes, HW promises
 - What the software writer needs to know to write and debug system/user programs
- Microarchitecture(**uarch**)/Organization
 - Specific implementation of an ISA
 - Not visible to the software
- Microprocessor
 - **ISA, uarch**, circuits
 - “Architecture” = ISA + microarchitecture



“Requirements Bottlenecks and Good Fortune Agents for Microprocessor Evolution”,
Y. Patt-IEEE 2001

1.1.2 计算机体系结构、组成和实现

- 体系结构的概念用于描述计算机系统设计的技术、方法和理论，包括以下三个方面：
 - (1) 计算机指令系统
 - (2) 计算机组成
 - (3) 计算机硬件实现
- 涵盖处理器和多处理器、存储器、输入/输出系统、互联与通信等计算机系统设计的主要内容
- 还涉及到性能评价、编译和操作系统技术
- 通过定量分析的途径，学习掌握现代计算机体系结构研究的基本方法

计算机组成

- 指令集结构的逻辑实现
 - 数据通路宽度
 - 专用功能部件的设置
 - 功能部件的并行性
 - 缓冲和排队技术
 - 预测技术
 - 可靠性技术
 - 控制机构的组成，等等

计算机的实现

- 处理器、主存的物理结构
- 器件的集成度和速度
- 信号传输
- 器件、模块、插件、底板的划分与连接
- 涉及的专用器件
- 电源、冷却
- 微组装技术
- 整机装配技术，等等



不同年代计算机体系结构研究的变化

| 年代 | 一些重要研究内容 | 典型计算机 |
|----------------|---------------------------------|---------------------------------|
| 1940年代 | 程序控制计算机、存储程序计算机 | ENIAC、EDVAC |
| 1960年代 | 指令系统 | IBM 360 系列机 |
| 1960年代 | 阵列机和并行处理 | ILLIAC IV |
| 1970年代 | 流水线、向量处理、微处理器 | Cray-1、Intel 4004 |
| 1980年代 | RISC、cache、流水线 | MIPS R1000、POWER |
| 1990年代 | SMP、CMP、指令级并行 | MIPS R10000、PowerPC 604 |
| 2000年以来 | SMT、功耗、Multi-core、Stream | Intel i7、Power 6、ARM、GPU |
| 2010年以来 | 功耗、异构计算、定制计算、指令扩展 | GPU、Intel AVX、RISC-V、TPU |

1.1 计算机体系结构的概念

1.1.1 计算机体系结构概念的演变

1.1.2 计算机体系结构、组成和实现

1.1.3 系列机和兼容



1.1.3 系列机和兼容

- 系列机（**family machine**）是具有相同体系结构，但组成和实现不同的一系列不同型号的计算机系统
 - **IBM**公司在推出**IBM S360**时首次提出的系列机的概念，被认为是计算机发展史上一个重要里程碑
 - 各计算机厂家仍按系列机研发产品
- 现代计算机不但系统系列化，其构成部件和软件也系列化
 - 如：微处理器（**CPU**）、硬盘、操作系统、高级语言等

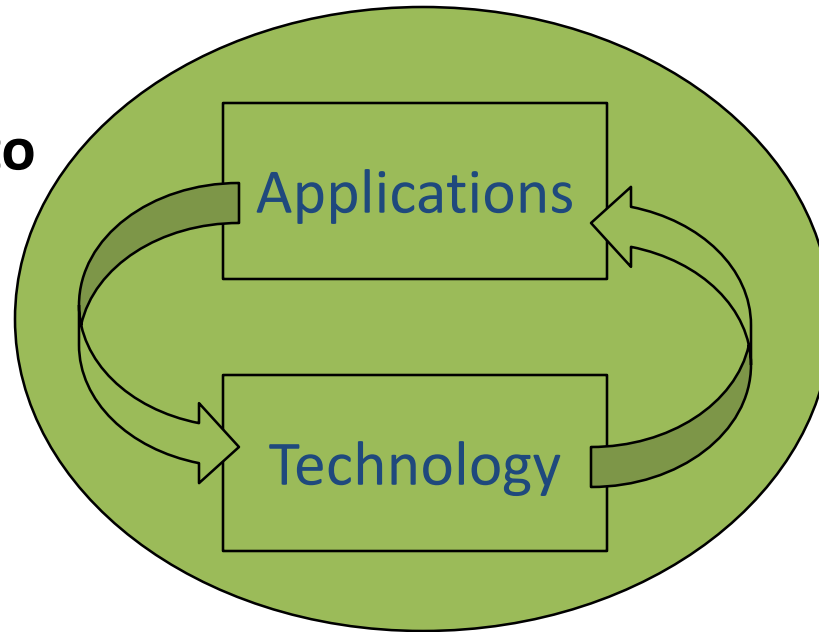
系列机(Family machine)

- 一种指令集结构可以有多种组成。同样，一种组成可以有多种物理实现。**系列机**就是指在**一个厂家**生产的具有相同的指令集结构，但具有不同组成和实现的一系列不同型号的机器。

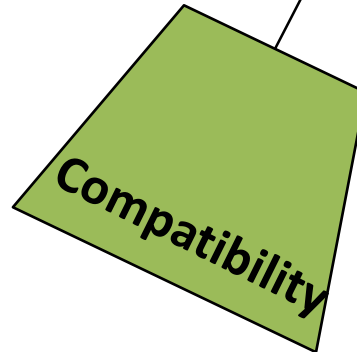
| 计算机 | 时间 | 处理器 | 字宽 | 主要I/O总线 |
|----------------|------|-------------|-----|----------------|
| PC和PC XT | 1981 | 8088 | 16位 | PC总线 |
| PC AT | 1982 | 80286 | 16位 | AT (ISA) |
| 80386 PC | 1985 | 80386 | 32位 | ISA/EISA |
| 80486 PC | 1989 | 80486 | 32位 | ISA+VL |
| Pentium PC | 1993 | Pentium | 32位 | ISA+PCI |
| Pentium II PC | 1997 | Pentium II | 32位 | ISA+PCI+AGP |
| Pentium III PC | 1999 | Pentium III | 32位 | PCI+AGP +USB |
| Pentium 4 PC | 2000 | Pentium 4 | 32位 | PCI-X+AGP +USB |

Architecture Continually Changing

Applications suggest how to improve technology, provide revenue to fund development



Improved technologies make new applications possible



Cost of software development makes compatibility a major force in market

软件兼容

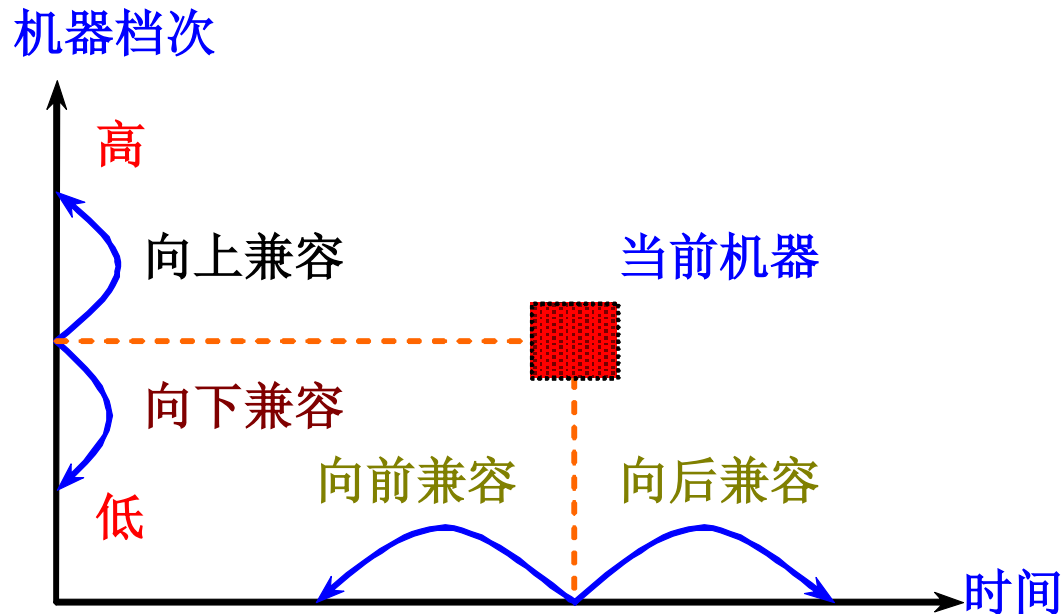
- 长期以来，程序员希望有一个稳定的软件环境，使他们编制出来的程序能够在更加广泛计算机类型中得到长期的应用。
- 软件兼容 (**software compatibility**)
 - 系列机具有相同的体系结构，软件可以在系列计算机的各档机器上运行。
 - 同一个软件可以不加修改地运行于体系结构相同的各档机器，而且它们所获得的结果一样，差别只在于有不同的运行时间。
 - 兼容分为二进制级兼容、汇编级兼容、高级语言兼容、数据级兼容等等。

兼容机

- 虽然程序员希望有一个稳定的软件环境，但是机器设计人员则希望根据硬件技术和器件技术的进展不断地推出新的机器。
- 兼容机（**compatible machine**）
 - 不同厂家生产的具有相同体系结构的计算机
 - 推动了部件标准的规范化，推进了计算机产品标准化的进程，降低生产和制造成本。
 - 在市场上有较强的竞争能力

兼容性

- 向上(下)兼容指的是按某档机器编制的程序，不加修改的就能运行于比它高(低)档的机器
- 向前(后)兼容指的是按某个时期投入市场的某种型号机器编制的程序，不加修改地就能运行于在它之前(后)投入市场的机器



Intel公司的x86系列

- Intel公司的x86系列微处理器在向后兼容方面是非常具有代表性的
- 1979年的8086到最新的Lake系列
- 由16位系统发展到64位系统
- 增加了保护方式指令集、MMX指令集和面向64位的扩展等
- 但它保持了极好的二进制代码级的向后兼容性

兼容对体系结构的影响

- 计算机系统及软件设计者的“障碍”
 - 系统软件的开发难度大
 - 需要保护巨大的应用软件宝库
- 向后兼容是才是软件兼容的根本特征，也是系列机的根本特征
 - 为了保证软件的兼容，要求指令集不改变，这无疑又妨碍计算机体系结构的发展
 - 向后兼容虽然削弱了系列机对体系结构发展的约束，但仍然是体系结发展的沉重包袱
 - 20世纪80年代具有RISC体系结构的微处理器在新结构、新技术应用等方面领先传统的CISC微处理器的主要原因之一

第1章 计算机系统量化分析基础

1.1 计算机体系结构的概念

1.2 计算机体系结构的发展

1.3 计算机系统设计和分析

1.2 计算机体系结构的发展

1.2.1 计算机分代和分型

1.2.2 软件的发展

1.2.3 应用的发展

1.2.4 相关核心技术的发展

1.2.5 体系结构的发展

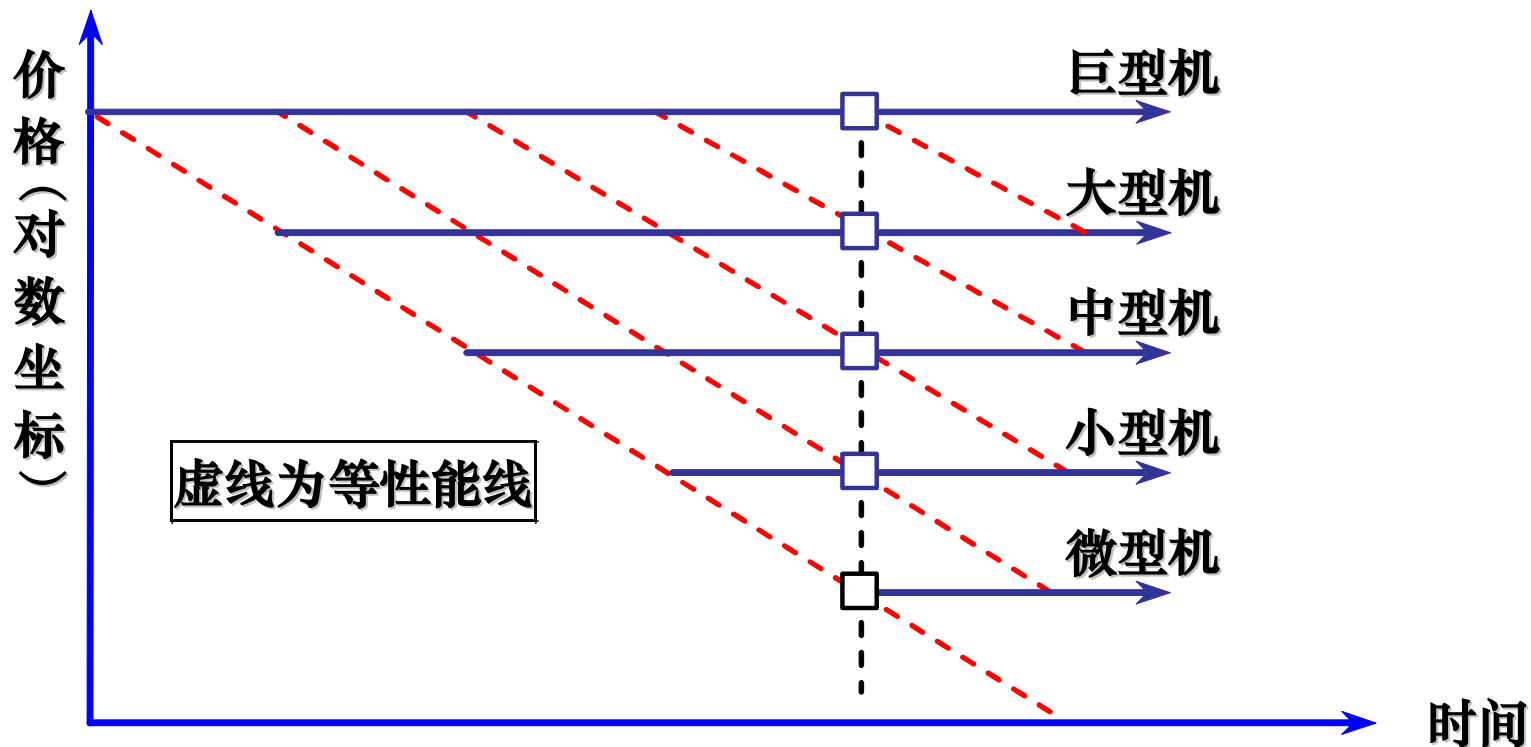
1.2.6 并行处理技术的发展

1.2.7 处理器芯片技术的发展

1.2.1 计算机的分代

| 分代 | 器件 | 体系结构技术 | 软件技术 | 典型机器 |
|--------------------|--------------------|-----------------------|-----------------|--|
| 第一代 (1945-1954) | 电子管和继电器 | 存储程序计算机、 程序控制I/O | 机器语言和汇编语言 | 普林斯顿ISA、 ENIAC IBM701 |
| 第二代 (1955-1964) | 晶体管、磁芯、印刷电路 | 浮点数据表示、寻址技术、中断、I/O处理机 | 高级语言和编译、批处理监控系统 | Univac LARC CDC1604 IBM7030 |
| 第三代 (1965-1974) | SSI和MSI、多层印刷电路、微程序 | 流水线、Cache、先行处理、系列计算机 | 多道程序和分时操作系统 | IBM360/370CDC 6600/7600、 DEC PDP-8 |
| 第四代 (1974-1990) | LSI和VLSI、半导体存储器 | 向量处理、分布式存储器 | 并行与分布处理 | Cray-1、IBM 3090、DEC VAX9000、 Convax-1 |
| 第五代 (1991-) | 高性能微处理器、大规模高密度电路 | 指令级并行、SMP、MP、MPP、网络 | 可扩展并行与分布处理 | SGI Cray T3E IBM xServer Sun E10000 |

1.2.1 计算机的分代



技术和性能的“下移”。新型体系结构的设计一方面是合理地增加计算机系统中硬件的功能比例。另一方面则是通过多种途径提高计算机体系结构中的并行性。

1. 2. 2 软件的发展

程序和数据使用的存储器容量不断增大。

1. 计算机语言与编译技术

2. 操作系统

3. 软件工具和SaaS (Software as a Service)

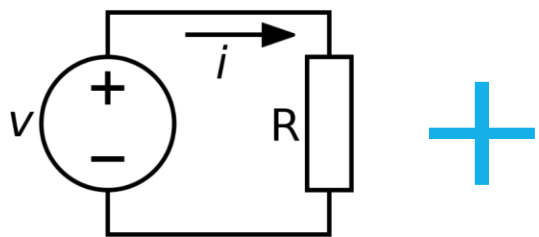
1.2.3 应用的发展

- 嵌入式计算机
包括移动设备和掌上计算机（关注功耗、成本等）
- 台式计算机
对性价比要求最高
- 服务器市场
可用性、高流量密度和可扩展性
- 数据中心
类似于服务器的要求，电能和散热系统很重要
- 巨型机（超级计算机）
面向浮点运算的数据中心计算机，内部网络需要更高的带宽和更低的延迟

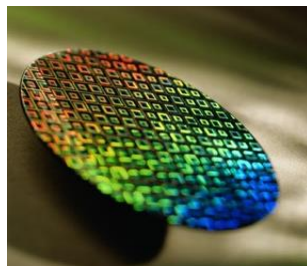
1.2.4 相关核心技术的发展

1. 集成电路逻辑技术

芯片(Chip)是一种把电路（如电源、电阻、电容、电感、二极管、晶体管、集成电路和电键等，构成的网络、硬件）小型化的方式，并制造在半导体晶圆表面上。



电路



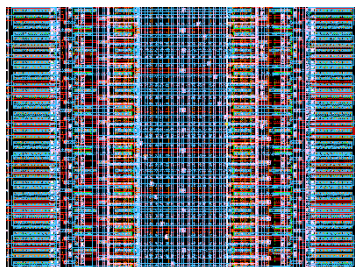
晶圆



芯片

1.2.4 相关核心技术的发展

1. 集成电路逻辑技术



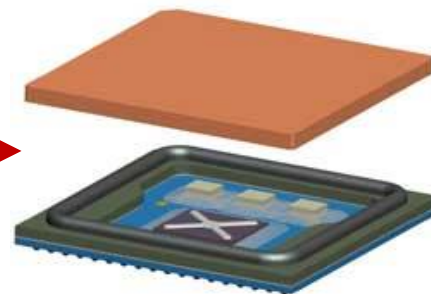
芯片设计

集成电路设计涉及对电子器件和器件间互连线模型的创建。



晶圆加工

晶圆加工是一系列化学处理步骤，使得电子电路逐渐形成在使用纯半导体材料制作的芯片上。

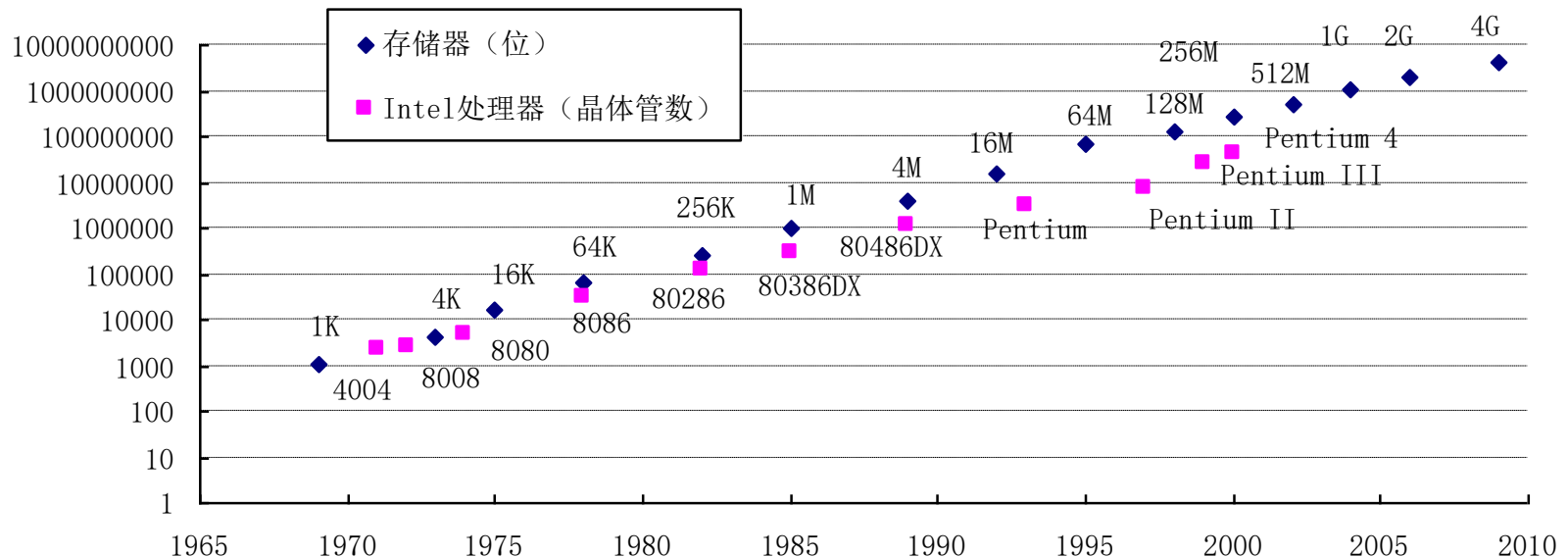


芯片封装测试

封装是将器件的核心晶粒封装在一个支撑物之内的过程，这个封装可以防止物理损坏以及化学腐蚀，并提供对外连接的引脚，之后将进行集成电路性能测试。

1.2.4 相关核心技术的发展

- 1965年，时任仙童公司研发实验室主任的摩尔（Gordon Mooer）在《Electronics》上撰文，认为集成电路密度大约每两年翻一番。
- 40年来，摩尔定律不但印证了集成电路技术的发展，也印证了计算机技术的发展。

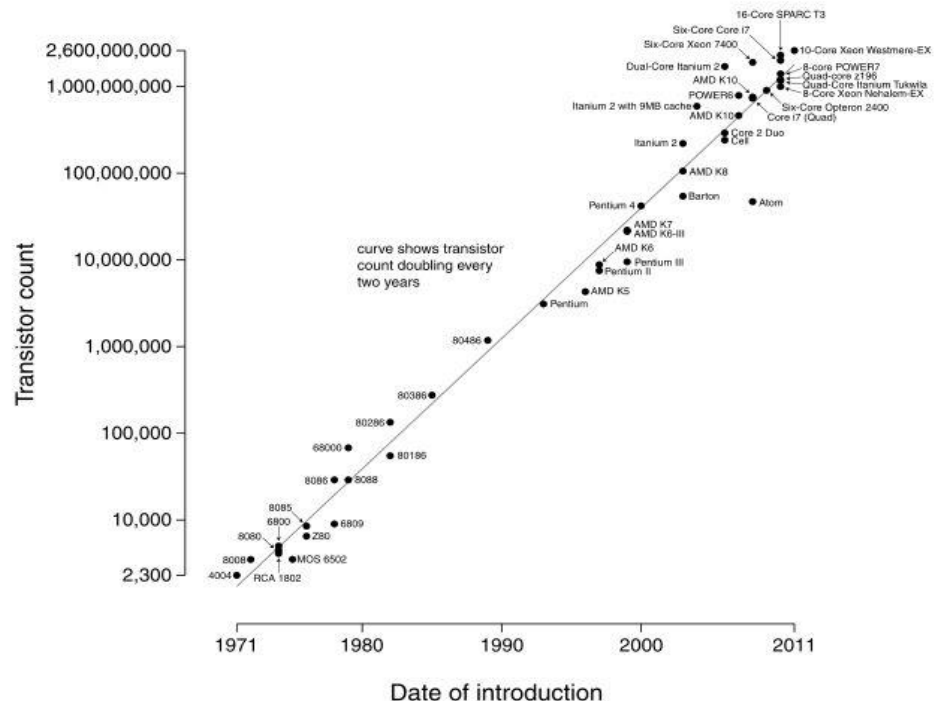


1.2.4 相关核心技术产品的发展

1. 集成电路逻辑技术（摩尔定律）

晶体管密度每年大约增加**35%**，约四年翻两番。晶片尺寸增长每年约**10-20%**。总之，一个芯片的晶体管数量每年增长**40-55%**，**18-24**个月翻一番

Microprocessor Transistor Counts 1971-2011 & Moore's Law



1.2.4 相关核心技术产品的发展

1. 集成电路逻辑技术（Dennard 缩放定律）

| 晶体管属性 ↴ | Dennardian 定律 ↴ | Post-Dennardian 定律 ↴ |
|----------------------|-----------------|----------------------|
| Δ 数量 Q ↴ | S^2 ↴ | S^2 ↴ |
| Δ 频率 F ↴ | S ↴ | S ↴ |
| Δ 电容 C ↴ | $1/S$ ↴ | $1/S$ ↴ |
| Δ 电压 V^2 ↴ | $1/S^2$ ↴ | 1 ↴ |
| =>Power = $QFCV^2$ ↴ | 1 ↴ | S^2 ↴ |
| =>利用率 = $1/Power$ ↴ | 1 ↴ | $1/S^2$ ↴ |

- Dennardian 1974年提出，晶体管尺寸每一代缩小30%，因此面积减小50%，电路减少30%的延迟，因此可以增加约1.4倍的工作频率。为了保持电场稳定，电压降低30%，能量降低65%，功率降低50%。
- 由于晶体管固有电压预充电压，当电压的下降与晶体管数量不能抵消，Power成指数增长（Post-Dennardian 定律）。