



计算机组成和体系结构 第一讲



四川大学 网络空间安全学院

版权声明

本文中所使用的图片、链接等使用归原版权人所有。

本文采用[创作共用署名-非商用使用-相同方式共享 4.0 国际版许可证](#)
[\(Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International, CC BY-NC-SA 4.0\)](#)
授权。



目录

00 / 课程介绍

01 / 绪论1：计算机组成和体系结构概述

02 / 绪论2：计算机发展历史

03 / 绪论3：计算机层次结构与云计算

04 / 绪论4：冯-诺依曼模型与并行计算



00

课程介绍





- 课程对象：四川大学网络空间安全专业二年级本科生
- 任课教师：
 - 课序号01：王启旭 (qixuwang@scu.edu.cn)
 - 助教：汪邓喆 (2078177077@qq.com)
 - 课序号02：高凯 (kaigao@scu.edu.cn)
 - 助教：章蕾 (zhanglei66@qq.com)
 - 课序号03：李贝贝 (libeibei@scu.edu.cn)
 - 助教：宋涵 (songhan@stu.scu.edu.cn)



- 内容丰富
 - 48学时
 - 大量的专业术语、
- 偏硬件
 - 涉及现代通用计算机各组成部分的工作原理、实现机制以及相互关系
 - 要求已经选修过数字电路设计等相关课程
- 与后续课程联系紧密
 - 学习内容涉及到操作系统、编译原理、高性能计算、计算机网络等



课程评分细则



四川大学
SICHUAN UNIVERSITY

最终成绩=出勤(20%)+ 作业(25%)+ 期中考试(15%)+ 期末考试(40%)

1. 课堂出勤占总成绩的20%，超过3次考勤未到则考勤记为**0分**；
2. 每周课程对应一个课程作业，课程作业占总成绩的25%；如发现作业抄袭则双方此次作业直接记为**0分**，两次抄袭则课程**不及格并上报学校**；
3. 在期中考试周，课程将进行期中考试（英文考卷，闭卷），考试范围为前半期学习的内容，占总成绩的15%；
4. 在期末考试周，课程将进行期末考试（英文考卷，闭卷），考试范围为后半期学习的内容，占总成绩的40%。
5. 考试作弊直接记为**不及格并上报学校**



- 基本目标：通过课程学习了解现代通用计算机各主要组成部分的硬件实现以及对应的实现逻辑

知其然

- 进阶目标：通过课程学习初步认识层次化抽象、指令集驱动硬件实现、性能-功耗-价格之间的权衡等系统设计思想

知其所以然



01

绪论 1

计算机组成和体系结构概述





计算机组成和体系结构的关系



四川大学
SICHUAN UNIVERSITY

计算机组成 (Organization)

- 计算机系统物理层面的各项组件以及相互之间的关联方式
- 例如：电子电路设计、控制信号、内存类型等等
- 通过学习计算机组成，了解**计算机是如何工作的**

计算机体系结构 (Architecture)

- 计算机系统实现逻辑层面的各项组件以及相互之间的关联
- 例如：指令集、指令编码、寻址模式等等
- 通过学习计算机体系结构，了解**如何设计计算机**

参考: Linda Null and Julia Lobur, *"The Essentials of Computer Organization and Architecture"*, 4th Edition

计算机组成和计算机体系结构密切相关：

- CPU设计：集成电路处理器 + 指令集
- 内存：不同类型的内存 + 层次化内存结构、段页式内存管理等



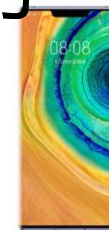
从服务器到移动设备，通用计算机的组成和体系结构是相似的



- CPU
 - Intel E5-2600v4
 - 2.4 GHz
 - 3级缓存15MB/25MB/35MB
- 内存
 - ECC DDR4 768G
- 存储
- 网卡
- PCIe、USB
- ...



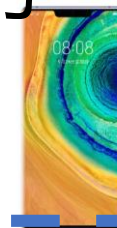
- CPU
 - Intel i7-8665U
 - 1.9 GHz
 - 缓存8MB
- 内存
 - DDR4 16G
- 存储
- 网卡
- 显卡
- PCIe、USB
- ...



- CPU
 - 麒麟990
 - 2.86 GHz
- 内存
 - 8G
- 存储
- 网卡
- 相机
- 耳机
- ...



从服务器到移动设备，通用计算机的组成和体系结构是相似的



- CPU

- Intel E5-2600v4
- 2.4 GHz
- 3级缓存15MB/25MB/35MB

- 内存

- ECC DDR4 768G

- 存储

- 网卡

- PCIe、USB

- ...

- CPU

- Intel i7-8665U
- 1.9 GHz
- 缓存8MB

- 内存

- DDR4 16G

- 存储

- 网卡

- CPU

- 麒麟990
- 2.86 GHz

- 内存

- 8G

- 存储

- 网卡

1. 解释并执行程序的中央处理器 (Central Processing Unit, CPU)



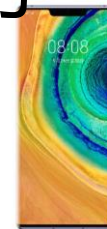
从服务器到移动设备，通用计算机的组成和体系结构是相似的



- CPU
 - Intel E5-2600v4
 - 2.4 GHz
 - 3级缓存15MB/25MB/35MB
- 内存
 - ECC DDR4 768G
- 存储
- 网卡
- PCIe、USB
- ...



- CPU
 - Intel i7-8665U
 - 1.9 GHz
 - 缓存8MB
- 内存
 - DDR4 16G
- 存储
- 网卡
- PCIe、USB
- ...



- CPU
 - 麒麟990
 - 2.86 GHz
- 内存
 - 8G
- 存储
- 网卡
- 耳机
- ...

2. 储存程序和数据内存



从服务器到移动设备，通用计算机的组成和体系结构是相似的



- CPU
 - Intel E5-2600v4
 - 2.4 GHz
 - 3级缓存15MB/25M

- 内存
 - ECC DDR4 768G

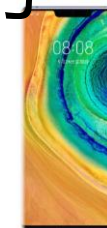
- 存储
- 网卡
- PCIe、USB
- ...



- CPU
 - Intel i7-8665U
 - 1.9 GHz

- 内存
 - DDR4 16G

- 存储
- 网卡
- 显卡
- PCIe、USB
- ...



- CPU
 - 麒麟990
 - 2.86 GHz

- 内存
 - 8G

- 存储
- 网卡
- 相机
- 耳机
- ...

3. 与外界设备进行数据交互的其它组件



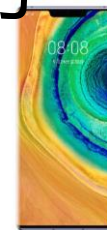
从服务器到移动设备，通用计算机的组成和体系结构是相似的



- CPU
 - Intel E5-2600v4
 - 2.4 GHz
 - 3级缓存15MB/25MB/35MB
- 内存
 - ECC DDR4 768G
- 存储
- 网卡
- PCIe、USB
- ...



- CPU
 - Intel i7-8665U
 - 1.9 GHz
 - 缓存8MB
- 内存
 - DDR4 16G
- 存储
- 网卡
- PCIe、USB
- ...



- CPU
 - 麒麟990
 - 2.86 GHz
- 内存
 - 8G
- 存储
- 网卡
- 相机
- 耳机
- ...

这些数字代表什么意思？



计算机系统的性能：常见计量单位



四川大学
SICHUAN UNIVERSITY

- 赫兹 (Hertz, Hz)
 - 物理意义：处理能力计量单位，单位时间时钟周期数
 - 常见对象：CPU、GPU等
 - 常见取值：GHz ($=10^9$ Hz)
- 字节 (Byte, B)
 - 物理意义：存储容量计量单位，存储介质可表示的字节数量
 - 常见对象：内存、硬盘等
 - 常见取值：GB ($=2^{20}$ Bytes), TB ($=2^{30}$ Bytes)
- 每秒传输比特/字节数 (Bit-per-second, bps或Byte-per-second, Bps)
 - 物理意义：传输速率计量单位，单位时间传输的比特数或字节数
 - 常见对象：网卡、PCI接口、USB接口等
 - 常见取值：Mbps ($=10^6$ bps), Gbps ($=10^9$ bps), GBps ($=\text{GB/s}=10^9$ Bps= 8×10^9 bps)
- 秒 (Second, s)
 - 物理意义：时间计量单位，某项操作所需时间
 - 常见对象：时钟周期、执行时间、传输延迟等
 - 常见取值：ms ($=10^{-3}$ s), ns ($=10^{-9}$ s) 等



计算机系统的性能：常见计量单位



四川大学
SICHUAN UNIVERSITY

- 赫兹 (Hertz, Hz)

- 物理意义：处理能力计量单位，单位时间时钟周期数
- 常见对象：CPU、GPU等
- 常见取值：GHz ($=10^9$ Hz)

$$f = \frac{1}{T}, \text{ 时钟频率是时钟周期的倒数}$$

- 字节 (Byte, B)

- 物理意义：存储容量计量单位，存储介质可表示的字节数量
- 常见对象：内存、硬盘等
- 常见取值：GB ($=2^{20}$ Bytes), TB ($=2^{30}$ Bytes)

$$r = \frac{D}{t}, \text{ 传输速率等于传输数据量除以时间}$$

- 每秒传输比特/字节数 (Bit-per-second, bps或Byte-per-second, Bps)

- 物理意义：传输速率计量单位，单位时间传输的比特数或字节数
- 常见对象：网卡、PCI接口、USB接口等
- 常见取值：Mbps ($=10^6$ bps), Gbps ($=10^9$ bps), GBps ($=\text{GB/s}=10^9$ Bps= 8×10^9 bps)

- 秒 (Second, s)

- 物理意义：时间计量单位，某项操作所需时间
- 常见对象：时钟周期、执行时间、传输延迟等
- 常见取值：ms ($=10^{-3}$ s), ns ($=10^{-9}$ s) 等



计算机系统的性能：常见计量单位



四川大学
SICHUAN UNIVERSITY

- 赫兹 (Hertz, Hz)
 - 物理意义：处理能力计量单位，单位时间时钟周期数
 - 常见对象：CPU、GPU等
 - 常见取值：GHz ($=10^9$ Hz)
- 字节 (Byte, B)
 - 物理意义：存储容量计量单位，
 - 常见对象：内存、硬盘等
 - 常见取值：GB ($=2^{20}$ Bytes), TB ($=2^{30}$ Bytes)
- 每秒传输比特/字节数 (Bit-per-second, bps或Byte-per-second, Bps)
 - 物理意义：传输速率计量单位，单位时间传输的比特数或字节数
 - 常见对象：网卡、PCI接口、USB接口等
 - 常见取值：Mbps ($=10^6$ bps), Gbps ($=10^9$ bps), GBps ($=\text{GB/s}=10^9$ Bps= 8×10^9 bps)
- 秒 (Second, s)
 - 物理意义：时间计量单位，某项操作所需时间
 - 常见对象：时钟周期、执行时间、传输延迟等
 - 常见取值：ms ($=10^{-3}$ s), ns ($=10^{-9}$ s) 等

后续课程中会介绍更多CPU性能指标



表示容量、速度等的前缀乘数

- Kilo- (K) = 1 thousand = 10^3 and 2^{10}
- Mega- (M) = 1 million = 10^6 and 2^{20}
- Giga- (G) = 1 billion = 10^9 and 2^{30}
- Tera- (T) = 1 trillion = 10^{12} and 2^{40}
- Peta- (P) = 1 quadrillion = 10^{15} and 2^{50}
- Exa- (E) = 1 quintillion = 10^{18} and 2^{60}
- Zetta- (Z) = 1 sextillion = 10^{21} and 2^{70}
- Yotta- (Y) = 1 septillion = 10^{24} and 2^{80}

表示时间、空间的前缀乘数

- Milli- (m) = 1 thousandth = 10^{-3}
- Micro- (μ) = 1 millionth = 10^{-6}
- Nano- (n) = 1 billionth = 10^{-9}
- Pico- (p) = 1 trillionth = 10^{-12}
- Femto- (f) = 1 quadrillionth = 10^{-15}
- Atto- (a) = 1 quintillionth = 10^{-18}
- Zepto- (z) = 1 sextillionth = 10^{-21}
- Yocto- (y) = 1 septillionth = 10^{-24}

1998年国际电子工程技术委员会 (International Electrotechnical Commission, IEC) 正式标准化二进制前缀乘数, “用以消除二进制和十进制可能存在的表示混淆”



02

绪论 2 计算机发展历史



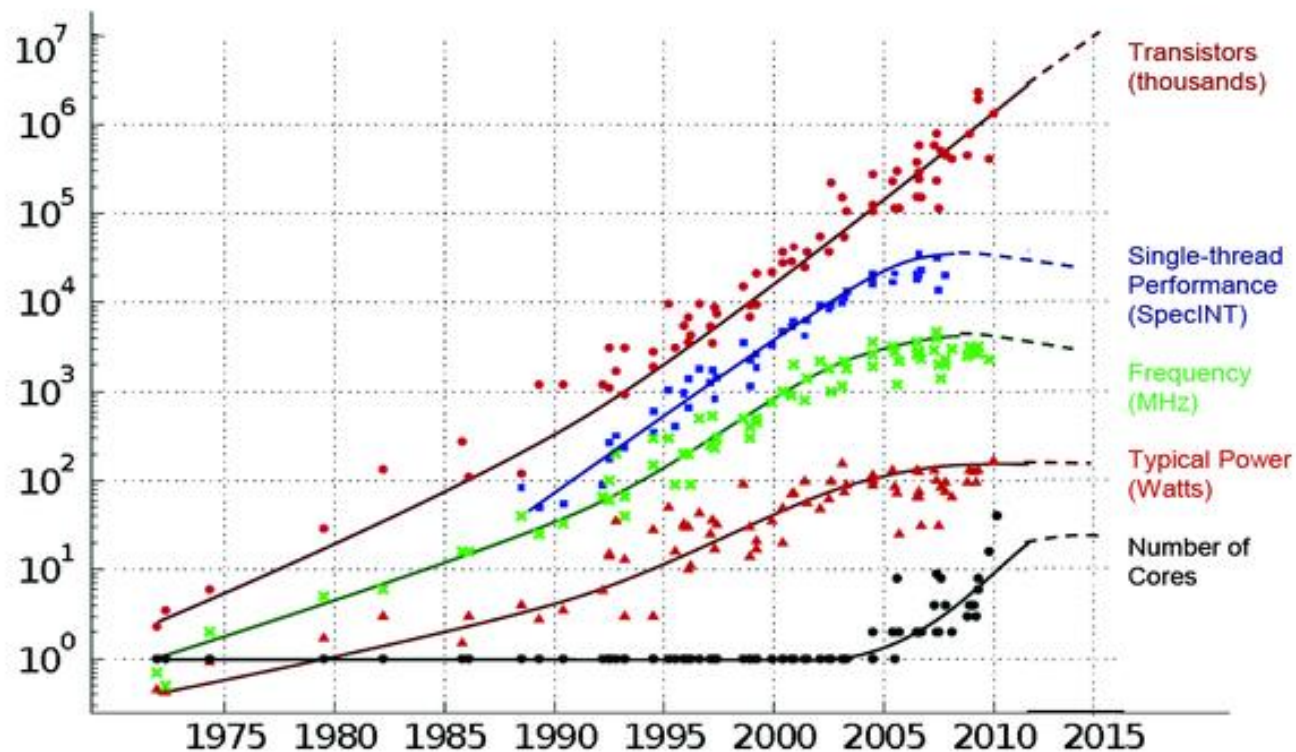


计算机发展：数据视角



四川大学
SICHUAN UNIVERSITY

35 YEARS OF MICROPROCESSOR TREND DATA



Original data collected and plotted by M. Horowitz, F. Labonte, O. Shacham, K. Olukotun, L. Hammond and C. Batten
Dotted line extrapolations by C. Moore

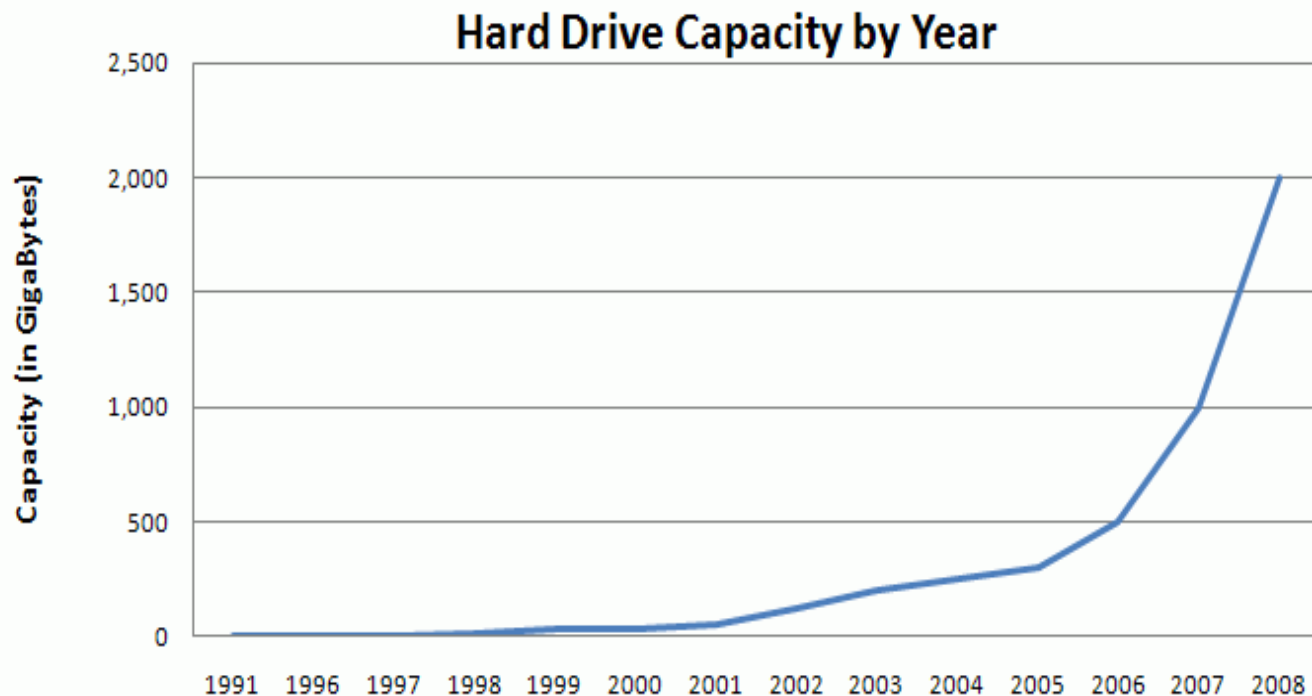
图片来源: Karl Rupp "40 Years of Microprocessor Trend Data"
<https://www.karlrupp.net/2015/06/40-years-of-microprocessor-trend-data/>



计算机发展：数据视角



四川大学
SICHUAN UNIVERSITY



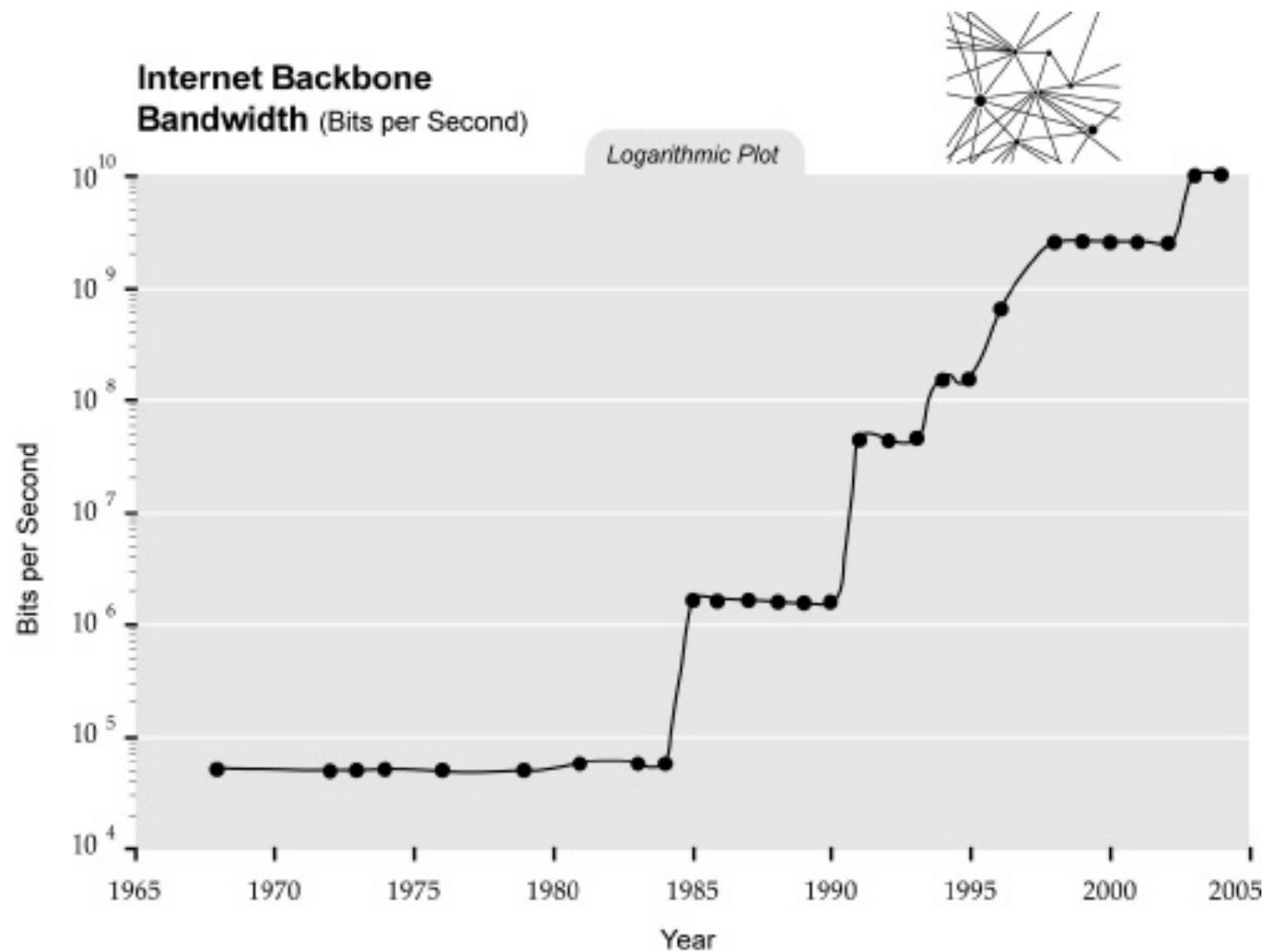
图片来源：Anton Shkel and Abraham Grair, UCLA CS 111, Lecture 10 “File System Performance”,
<http://web.cs.ucla.edu/classes/winter13/cs111/scribe/10c/>



计算机发展：数据视角



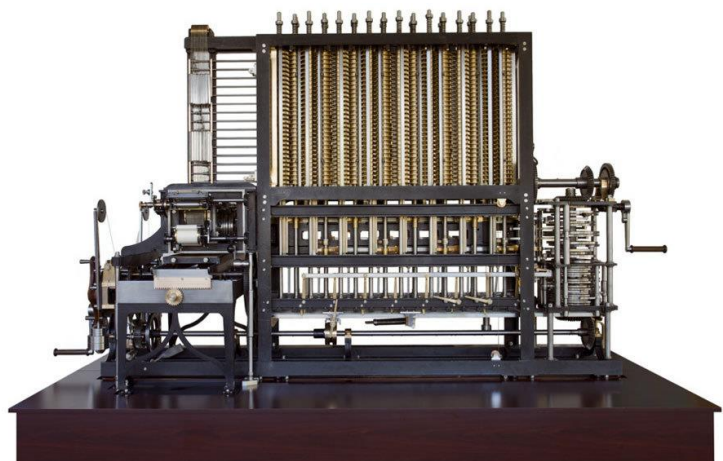
四川大学
SICHUAN UNIVERSITY



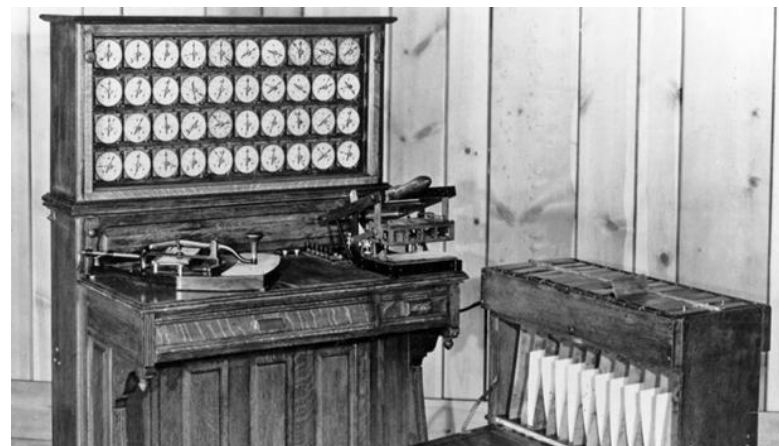
图片来源：
<http://www.singularity.com/charts/page81.html>



- 通常来说，我们讨论计算机发展历史的时候，主要是按照**计算单元**的技术飞跃将现代通用计算机分为四代
- 在现代计算机之前，已经出现了利用机械进行计算的工具



巴比奇差分机 (Differential Machine by Charles Babbage) 图片来源:
<https://www.npr.org/templates/story/story.php?storyId=121206408>



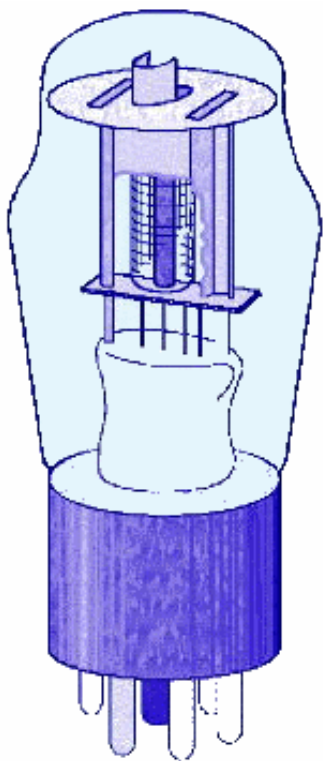
穿孔卡片制表机 (Punched Card Tabulator by Herman Hollerith) 图片来源:
<https://www.ibm.com/ibm/history/ibm100/us/en/icons/tabulator/>



第一代计算机：电子管（1945-1953）



四川大學
SICHUAN UNIVERSITY



- 阿塔纳索夫-贝瑞计算机（Atanasoff Berry Computer, 1937-1938）
 - 发明者：爱荷华州立大学的John Atanasoff 和 Clifford Berry
 - 世界上第一台电子计算机
 - 仅用于求解线性方程组
- 电子数字积分计算机（Electronic Numerical Integrator and Computer, ENIAC, 1946）
 - 发明者：宾夕法尼亚大学大学John Mauchly 和 J. Presper Eckert
 - 世界上第一台**通用（general-purpose）**计算机
- 103机（中国第一台电子计算机，1958）
- 同时间还有其它电子管计算机问世，请查阅参考书

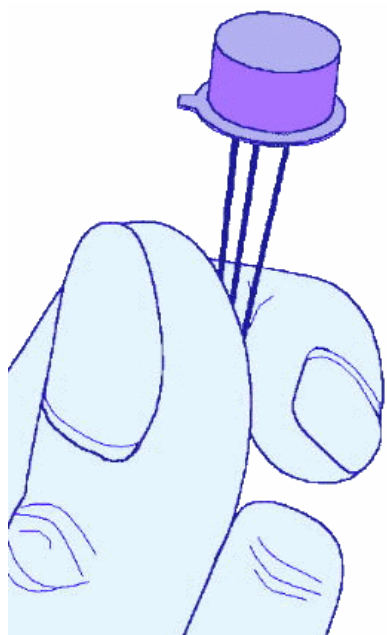
图片来源：Linda Null and Julia Lobur “The Essentials of Computer Organization and Architecture”, 4th Edition



第二代计算机：晶体管（1954-1965）



四川大學
SICHUAN UNIVERSITY



- TRADIC（贝尔实验室，1955）
 - 世界上第一台晶体管电子计算机
- IBM 7094（科学计算）以及1401（商用）
- ...
- 441B（中国第一台晶体管电子计算机，1964）

图片来源：Linda Null and Julia Lobur “The Essentials of Computer Organization and Architecture”, 4th Edition

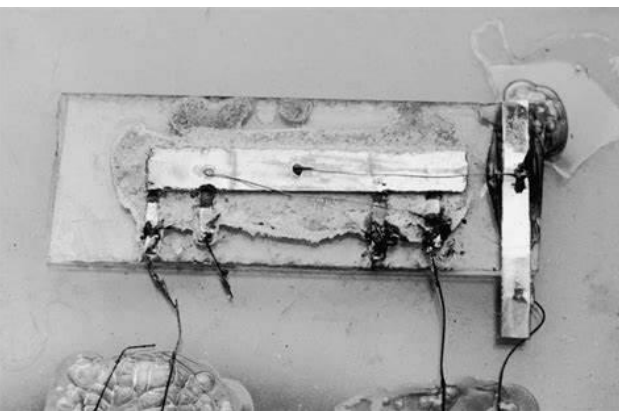


第三代计算机：集成电路（1965-1980）



四川大學
SICHUAN UNIVERSITY

- IBM System/360 (IBM, 1964)
 - 首次采用集成电路
- ...



图片来源：

<https://www.wsj.com/articles/the-chip-that-changed-the-world-1535311622>



第四代计算机：超大规模集成电路（1980-?）



四川大學
SICHUAN UNIVERSITY



- 超大规模集成电路（Very Large-scale Integrated Circuits, VLSI）的出现使得微处理器的发明成为可能
- 按照晶体管集成度分类：
 - 小规模集成（Small-scale Integration, SSI）： $10 \sim 10^2$
 - 中等规模集成（Medium-scale Integration, MSI）： $10^2 \sim 10^3$
 - 大规模集成（Large-scale Integration, LSI）： $10^3 \sim 10^4$
 - 超大规模集成电路： $10^4 \sim 10^6$ （2010年）
- Intel 4004：世界上第一块微处理器

图片来源：
https://de.wikipedia.org/wiki/VLSI_Technology



芯片发展两个著名定律



四川大学
SICHUAN UNIVERSITY

摩尔定律 (Moore' s Law)

- 1965年由Intel创始人戈登-摩尔 (Gordon Moore) 提出
- 芯片密度每18个月增长一倍

罗克定律 (Rock' s Law)

- 1960年代由Intel财务专家阿瑟-罗克 (Arthur Rock) 提出
- 芯片制造的成本每4年增长一倍



03

绪论 3

计算机层次化结构与云计算





计算机层次化结构



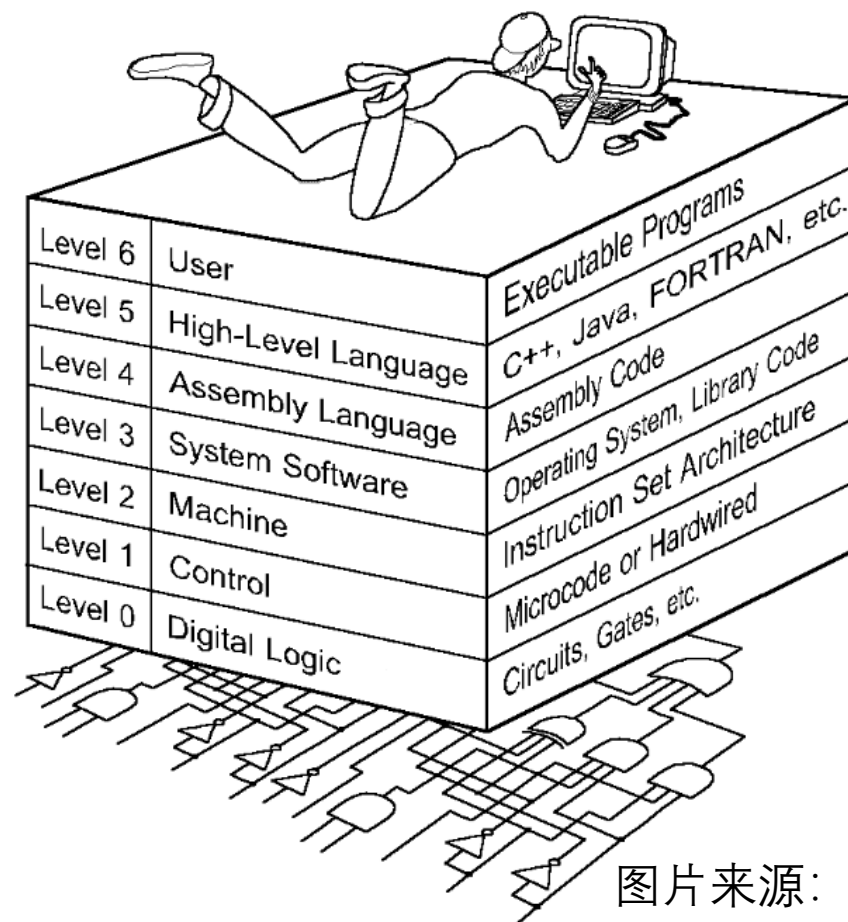
四川大学
SICHUAN UNIVERSITY

现代计算机是一个非常复杂的软硬件系统

- 程序员采用高级编程语言甚至是人机交互界面表达计算需求
- 底层实现依靠的是数字电路连接和数字信号处理

分解复杂性的一个重要方法是通过层次化抽象

- 每一层可以看做一个“虚拟机” (Virtual Machine), 具有特定的指令, 并且通过调用底层的“虚拟机”指令来执行特定的控制逻辑
- 其它例子: 网络协议栈, 软件工具栈



图片来源: Linda Null and Julia Lobur "The Essentials of Computer Organization and Architecture", 4th Edition



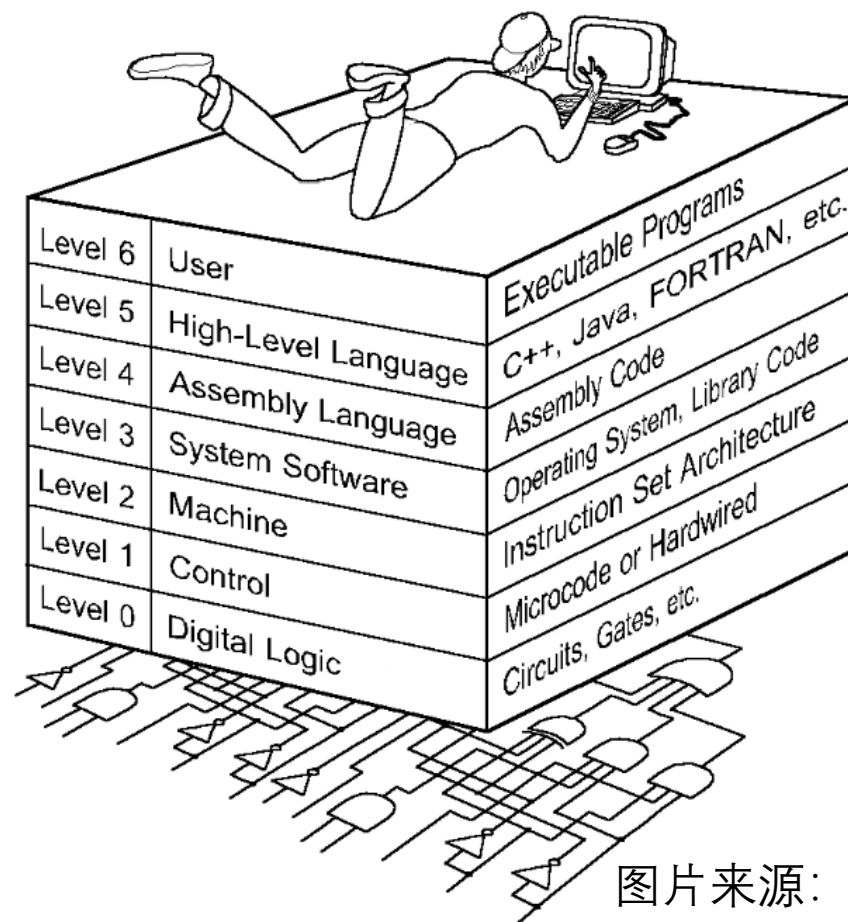
计算机层次化结构



四川大学
SICHUAN UNIVERSITY

Level 6: 用户层 (User)

- 直接面向终端用户的应用程序和交互界面
- 例子：桌面环境，浏览器，Office，IDE，游戏等



图片来源：Linda Null and Julia Lobur "The Essentials of Computer Organization and Architecture", 4th Edition



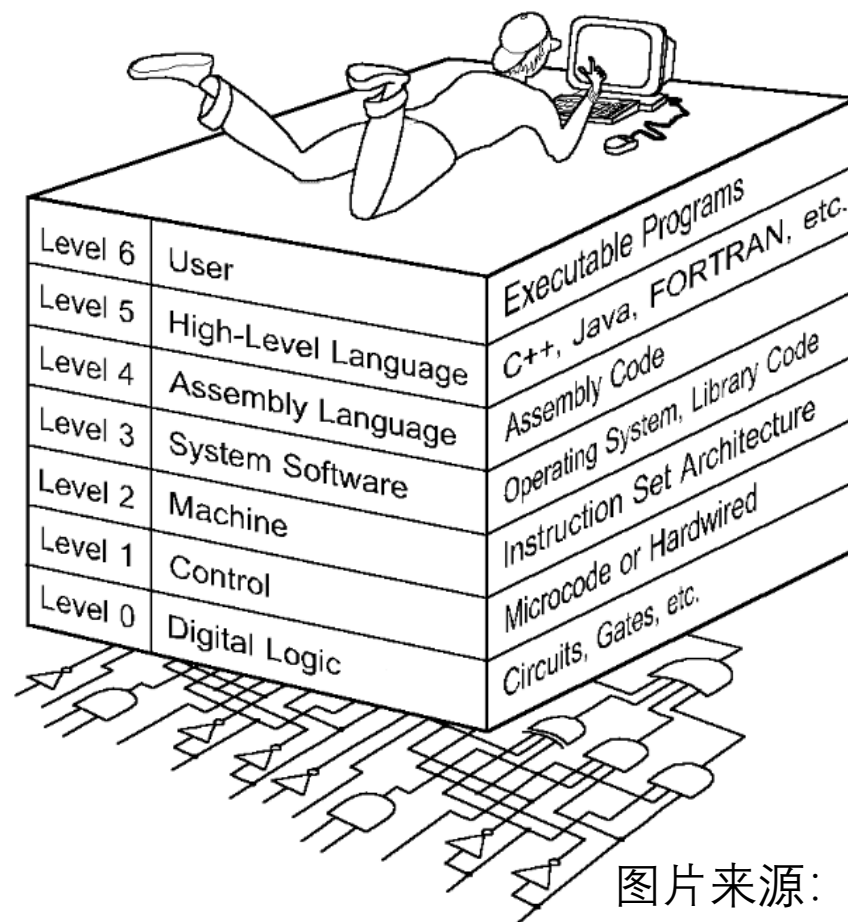
计算机层次化结构



四川大学
SICHUAN UNIVERSITY

Level 5: 高级编程语言 (High-level Language)

- 提供了数据类型、控制逻辑、内存管理等抽象的编程语言
- 例子: C/C++/C#, Pascal, Python, Java/Scala/Groovy, Javascript/Typescript, Haskell/F#等等.



图片来源: Linda Null and Julia Lobur "The Essentials of Computer Organization and Architecture", 4th Edition



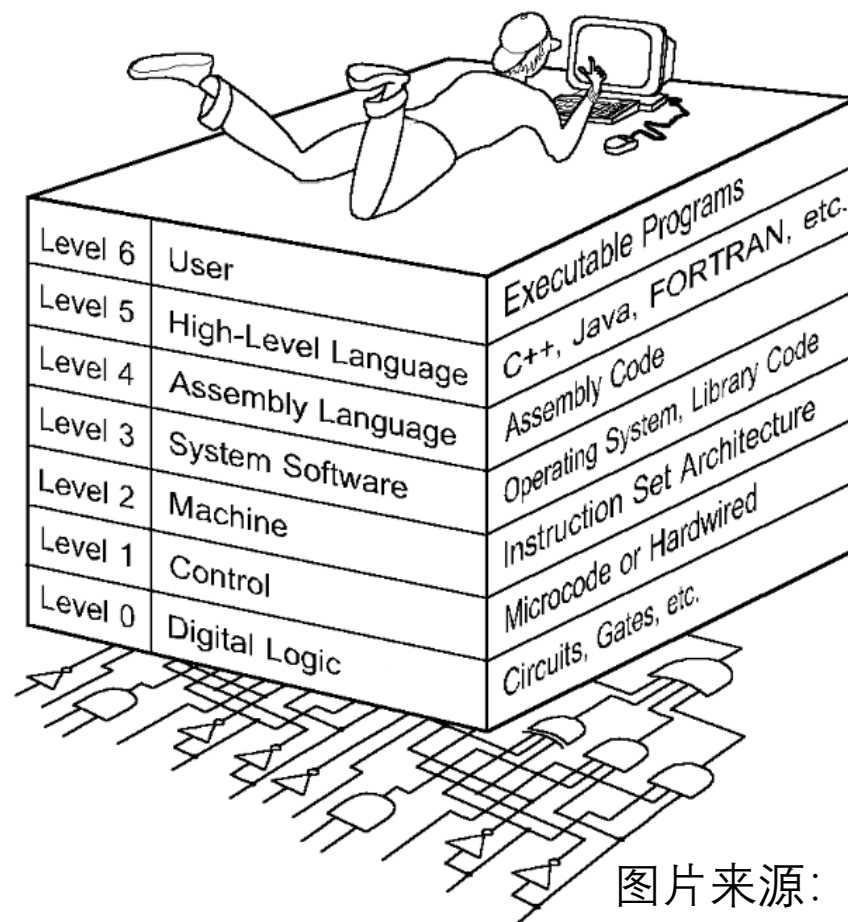
计算机层次化结构



四川大学
SICHUAN UNIVERSITY

Level 4: 汇编语言 (Assembly)

- 对底层硬件指令的直接翻译
- 例子: x86 assembly



图片来源: Linda Null and Julia Lobur "The Essentials of Computer Organization and Architecture", 4th Edition



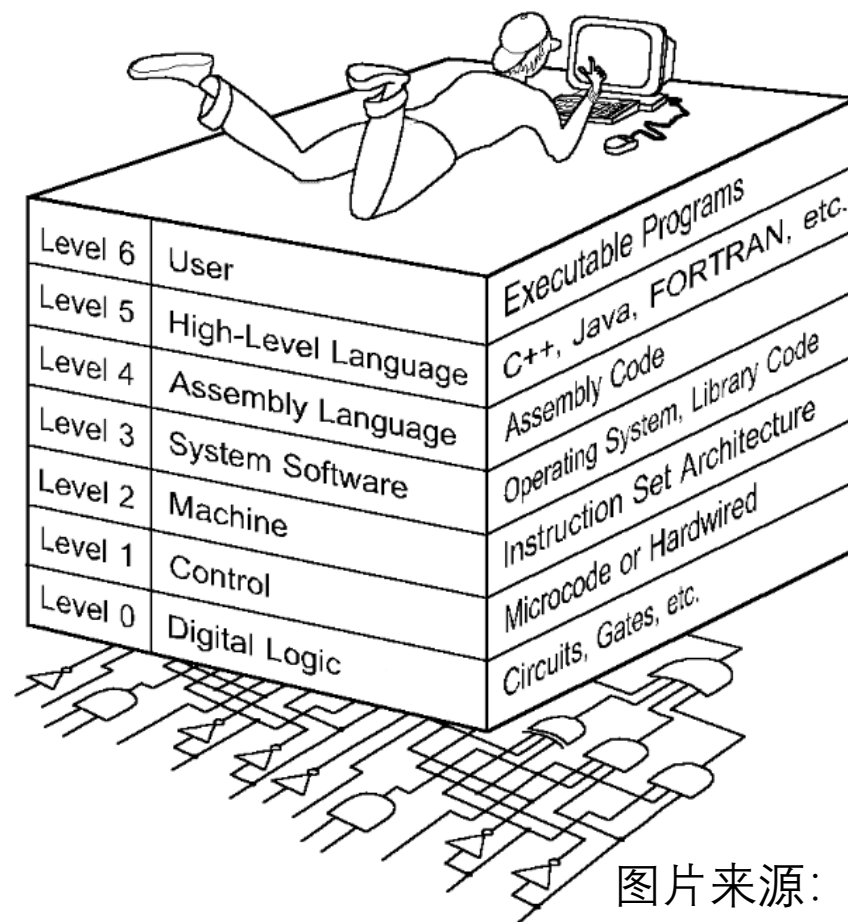
计算机层次化结构



四川大学
SICHUAN UNIVERSITY

Level 3: 系统软件 (System Software)

- 对系统资源提供调度和保护
- 例子: Linux, Windows, Unix, MacOS等



图片来源: Linda Null and Julia Lobur "The Essentials of Computer Organization and Architecture", 4th Edition



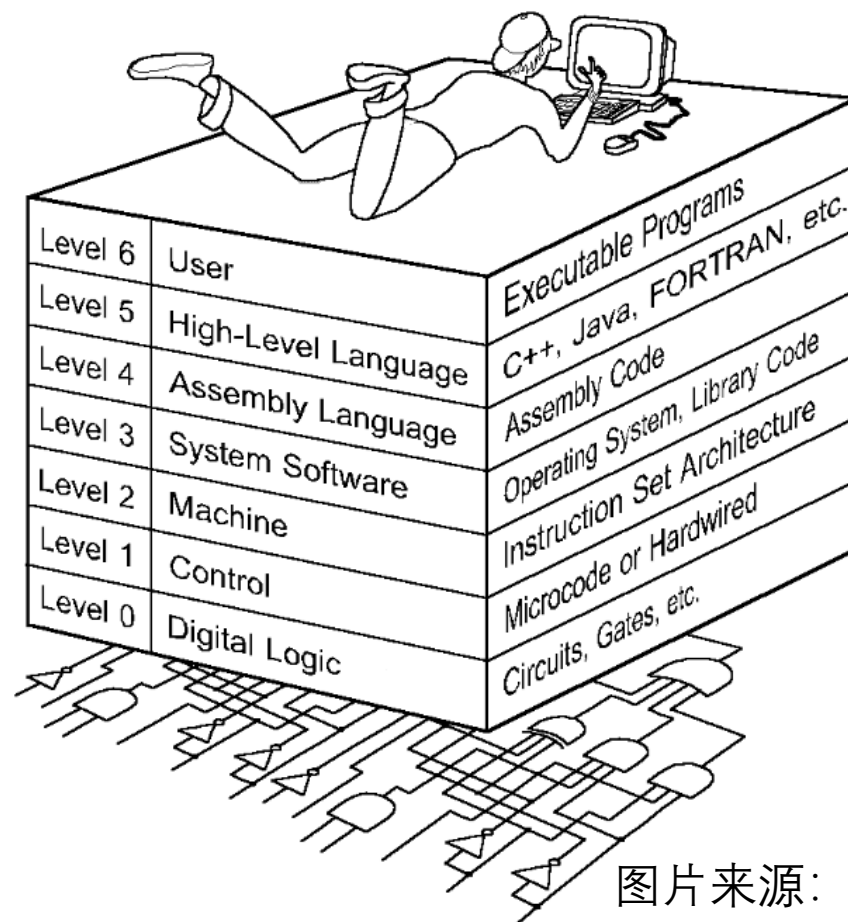
计算机层次化结构



四川大学
SICHUAN UNIVERSITY

Level 2: 机器层/指令集架构层
(Machine level/Instruction Set Architecture Level)

- 包含了底层芯片对应的指令集
- 不需要编译器、解释器或汇编器处理即可直接执行
- 例子: x86, ARM, RISC等



图片来源: Linda Null and Julia Lobur "The Essentials of Computer Organization and Architecture", 4th Edition



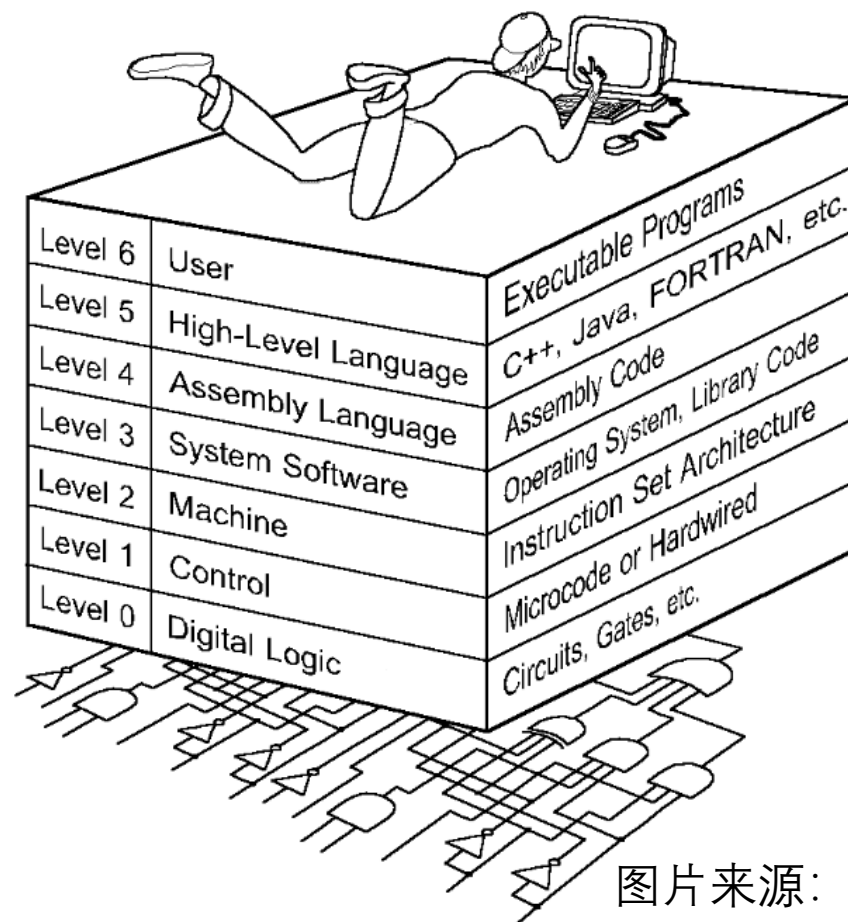
计算机层次化结构



四川大学
SICHUAN UNIVERSITY

Level 1: 控制层 (Control Level)

- 处理器通过控制单元进行指令的解码和执行以及进行数据的移动
- 控制单元可以是硬件直连 (hardwired) 或者可以被“微编程” (Microprogrammed)
- 例子: Intel microcode



图片来源: Linda Null and Julia Lobur "The Essentials of Computer Organization and Architecture", 4th Edition



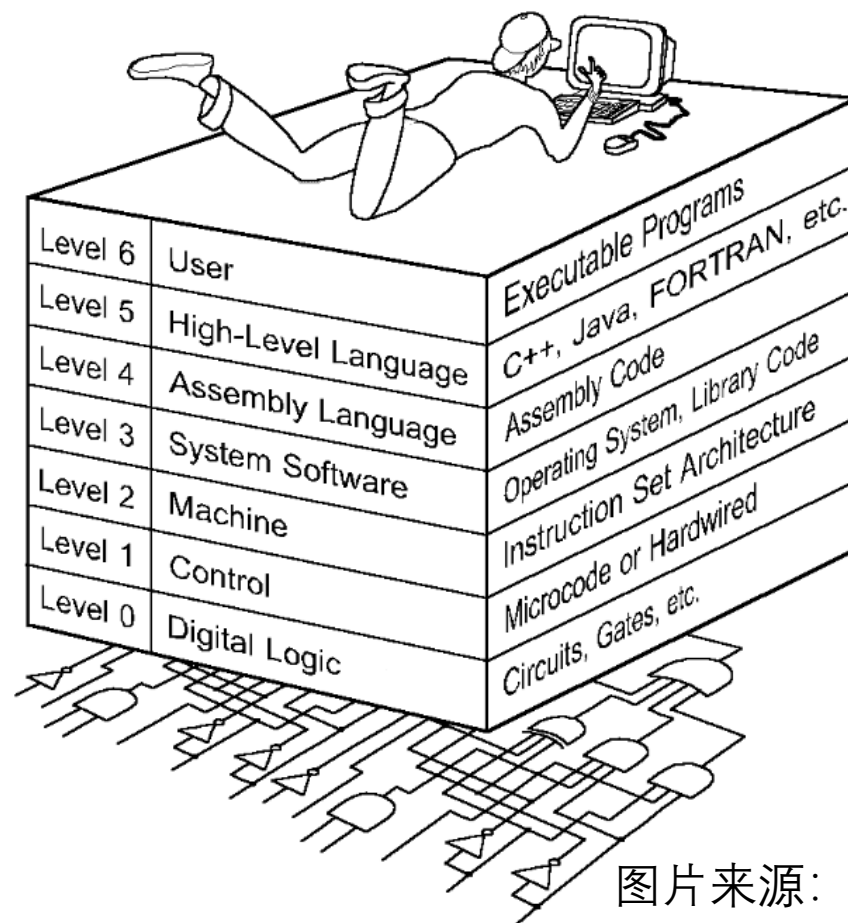
计算机层次化结构



四川大学
SICHUAN UNIVERSITY

Level 0: 数字逻辑层 (Digital Logic Level)

- 由门电路构成的数字电路
- 最终实现了上层所指定的功能



图片来源: Linda Null and Julia Lobur "The Essentials of Computer Organization and Architecture", 4th Edition



- 云计算出现的契机
 - 网络技术的发展
 - 大型数据中心的兴起：冗余计算资源
- 云计算相比购买硬件计算设备的优势
 - 硬件成本、维护成本、管理成本大幅下降
 - 服务弹性 (Elasticity)：对于云计算的用户来说，可根据实时的计算需求快速请求相应的计算资源
 - 资源利用率 (Utilization)：对于云计算的服务提供商来说，多租户共享资源，减少了计算资源闲置，提高了整体的利用率



Compute-as-a-service (计算即服务)

- Software-as-a-service (软件即服务)

- 例子: LAMP/Office 365/Steam/...

- Platform-as-a-service (平台即服务)

- 例子: Google App Engine/Azure Cloud Service/WeChat小程序等

- Infrastructure-as-a-service (基础设施即服务)

- 例子: Amazon EC2

- Function-as-a-service (函数即服务)

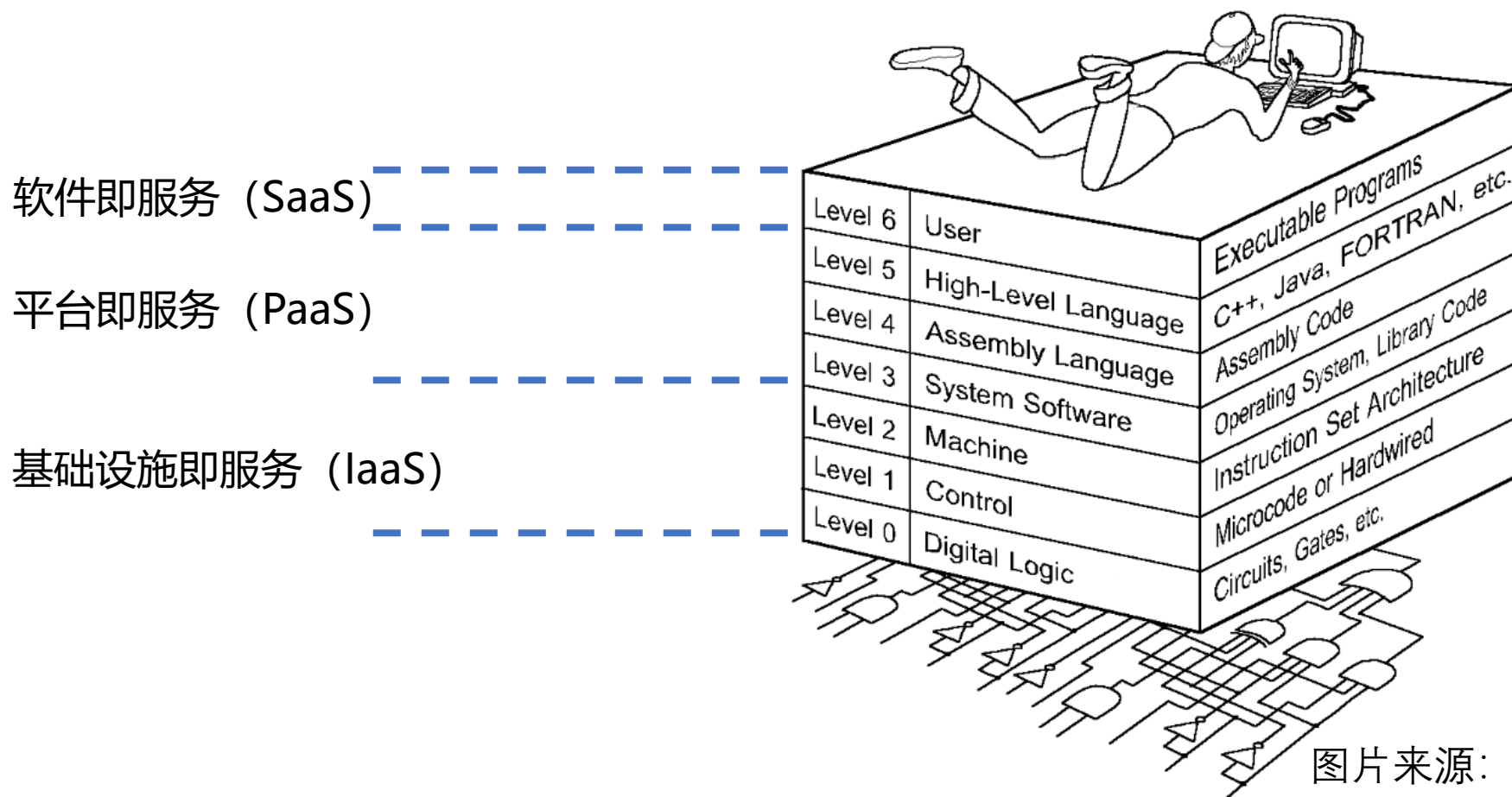
- 例子: Google Cloud Function、AWS Lambda



云计算与计算机层次化结构的关系



四川大学
SICHUAN UNIVERSITY



图片来源: Linda Null and Julia Lobur "The Essentials of Computer Organization and Architecture", 4th Edition



04

绪论 4

冯-诺依曼模型与并行计算





冯-诺依曼模型



四川大学
SICHUAN UNIVERSITY

- 首先由Mauchly和Eckert在ENIAC项目后提出，由冯-诺依曼最初公开发表
 - 改变ENIAC的功能需要手动对电子管进行连线
 - 每个特定的硬件配置（连线方式）只能解决一个特定的问题
- 核心思想：在内存中存储指令





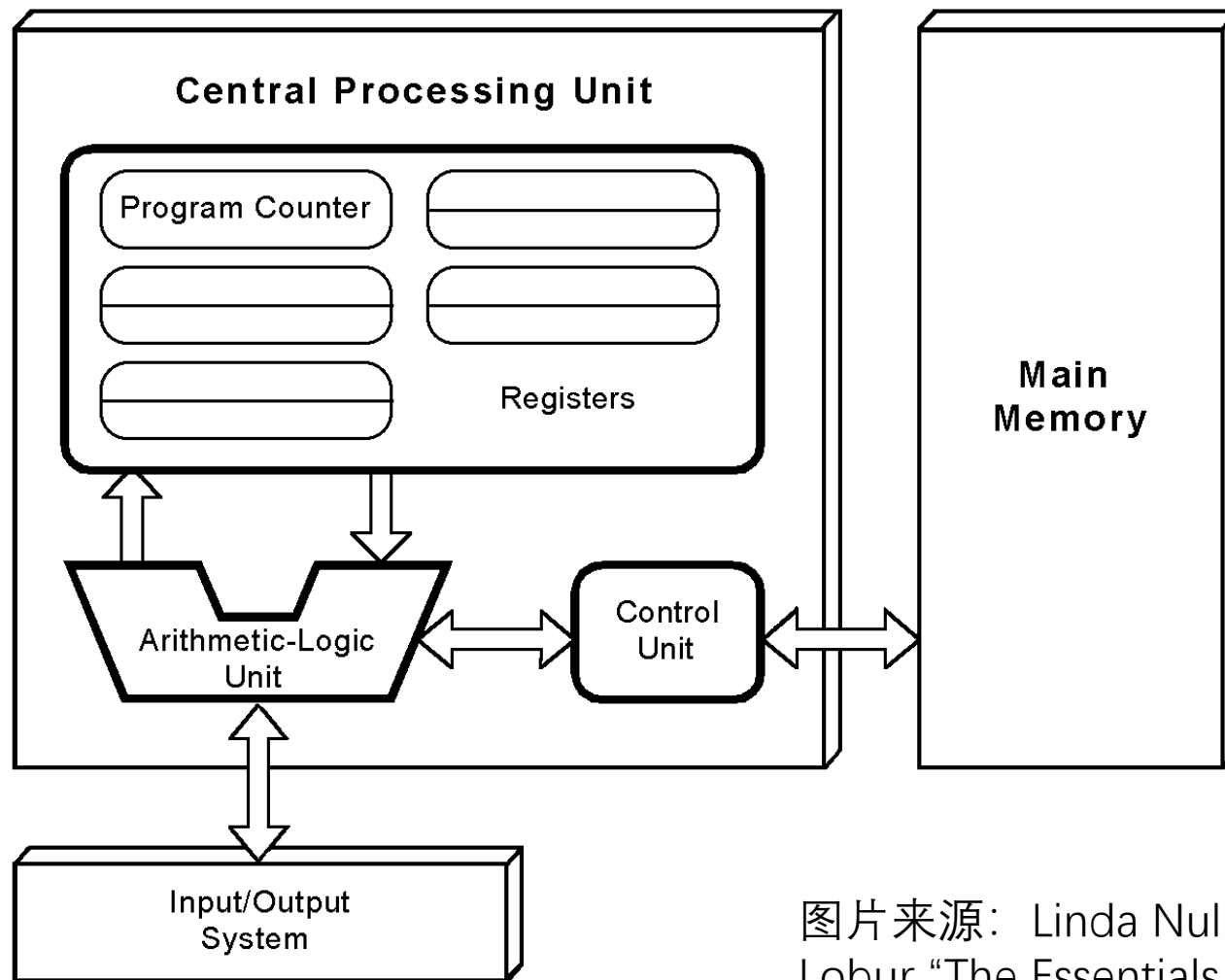
现在的通用计算机通常具有下面一些特征：

- 由三部分硬件系统构成
 - 中央处理单元（CPU）
 - 主存
 - 输入/输出系统
- 可以执行指令序列
- 在CPU和主存之间有且仅有一个数据通路（也被称为冯-诺依曼瓶颈）



图中是一个冯-诺依曼模型的计算机系统

程序的执行过程称为“冯-诺依曼执行周期”，也被称为“取-译码-执行”（Fetch-Decode-Execute）周期

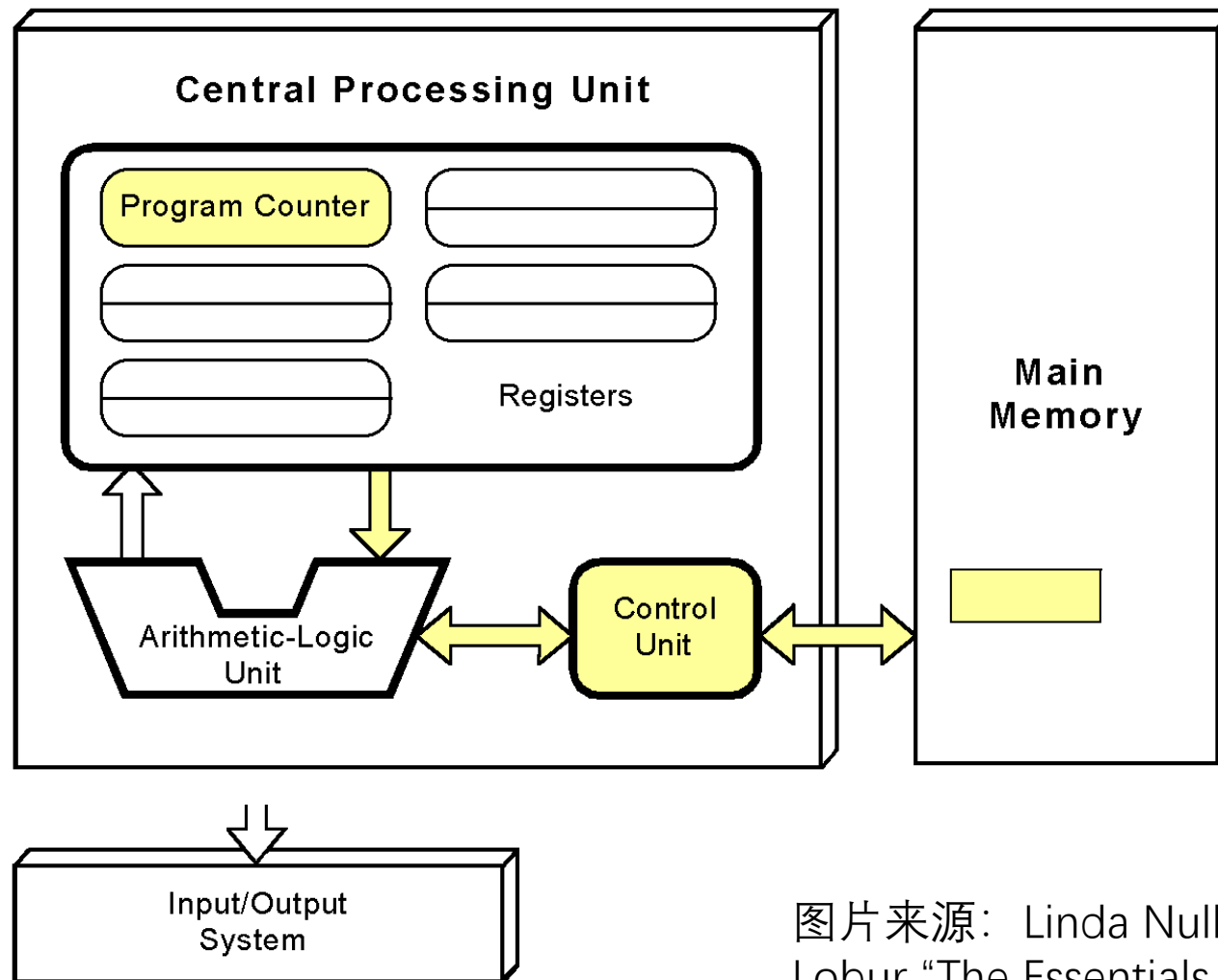


图片来源: Linda Null and Julia Lobur "The Essentials of Computer Organization and Architecture", 4th Edition



取 (Fetch)

CPU的控制单元根据程序计数器 (Program Counter, PC) 从主存中读取下一条需要执行的指令

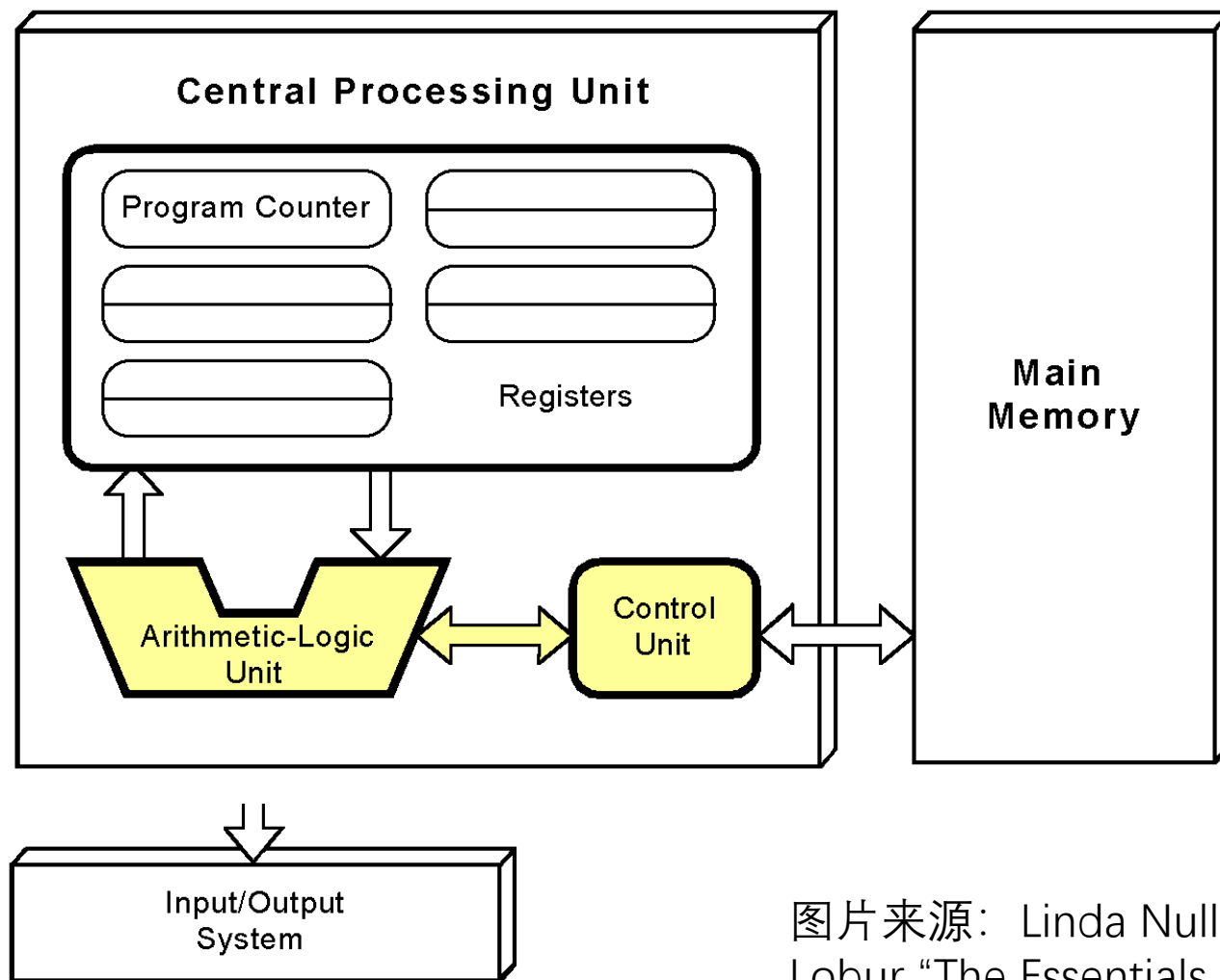


图片来源: Linda Null and Julia Lobur "The Essentials of Computer Organization and Architecture", 4th Edition



译码 (Decode)

获取的指令被译码成为算术逻辑单元 (Arithmetic-Logic Unit, ALU) 可以理解的形式



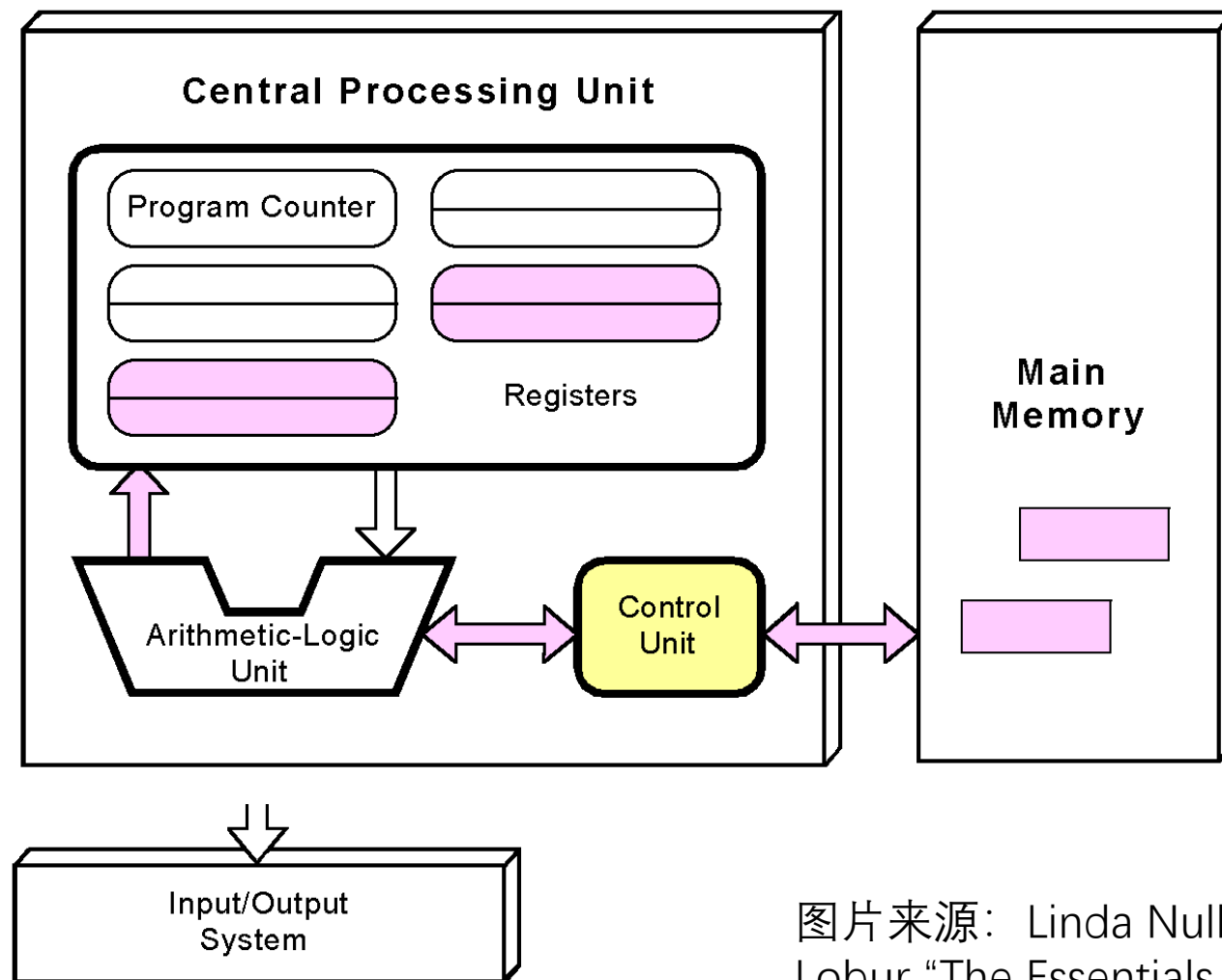
图片来源: Linda Null and Julia Lobur "The Essentials of Computer Organization and Architecture", 4th Edition



执行 (Execute)

根据译码的结果，CPU控制单元从主存中读取操作数对应的数据，并存入适当的寄存器 (register) 中

为什么这一步不是“取”？



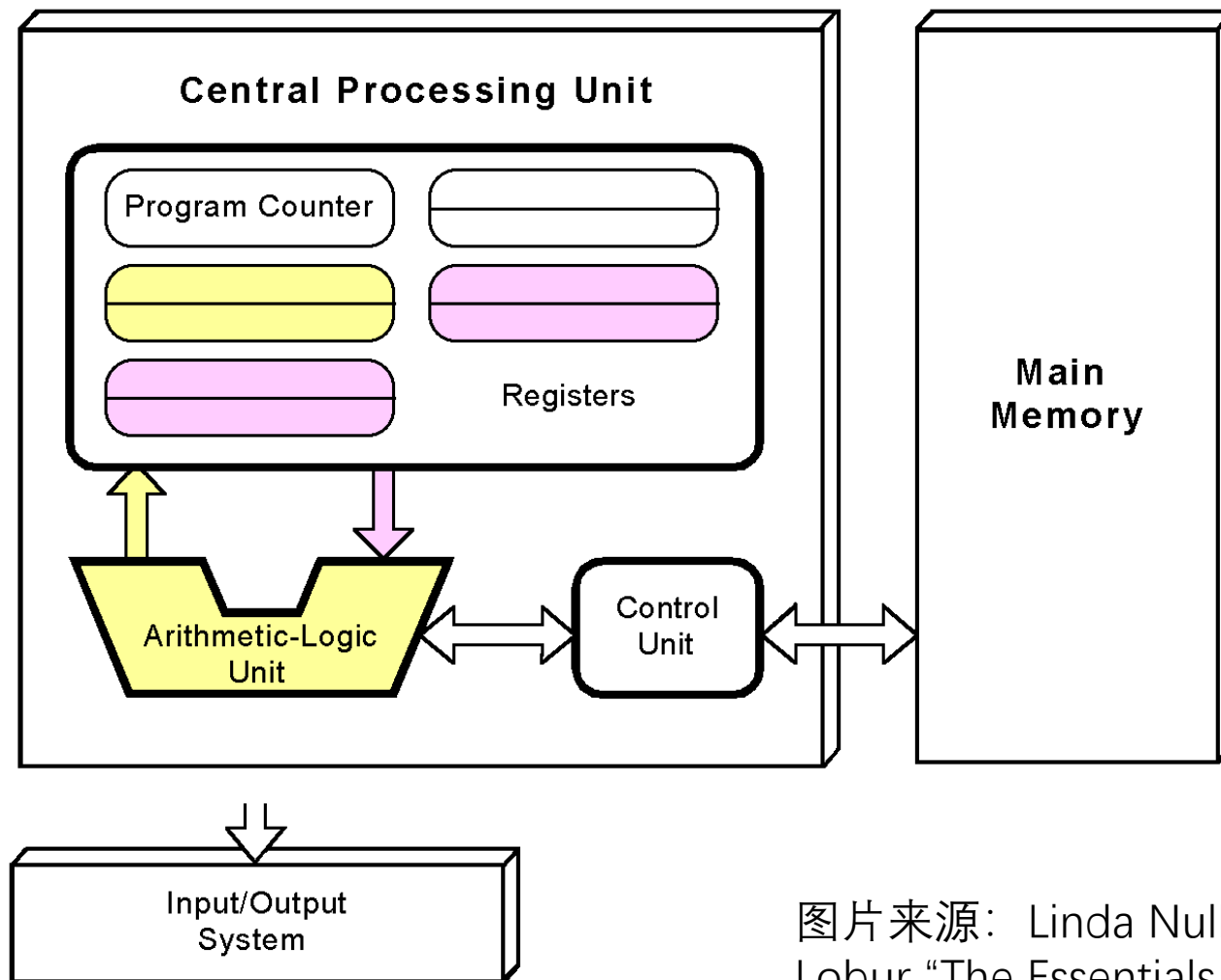
图片来源: Linda Null and Julia Lobur "The Essentials of Computer Organization and Architecture", 4th Edition



执行 (Execute)

根据译码的结果，CPU控制单元从主存中读取操作数对应的数据，并存入适当的寄存器中

ALU执行指令之后，将结果写回寄存器或主存



图片来源: Linda Null and Julia Lobur "The Essentials of Computer Organization and Architecture", 4th Edition



总结



四川大学
SICHUAN UNIVERSITY

- 认识计算机系统
 - 主要的系统组件
 - 相关性能指标以及符号表示
- 现代通用计算机发展历史
 - 四代计算机采用的核心技术
 - 电子管、晶体管、集成电路、超大规模集成电路
 - 每代计算机的标志性产品和时间
 - 摩尔定律和罗克定律
- 现代计算机层次化结构
 - 用户层、高级语言层、汇编语言层、系统软件层、机器层、控制层、数字电路层
 - 云计算及常见云计算类型
- 冯诺依曼模型
 - 冯-诺依曼模型的特征
 - 取译码执行的流程



四川大學
SICHUAN UNIVERSITY

谢谢！