面向摩尔定律的设计

使用抽象简化设计

加速大概率事件

通过并行提供性能

通过流水线提供性能

通过预测提供性能

存储器层次

通过冗余提供可靠性

执行时间：响应时间，计算机完成某个任务所需的总时间

吞吐率：带宽，单位时间完成的任务数

性能=1/执行时间

运行时间：总的响应时间，包括处理, I/O操作, OS 开销, 空闲时间

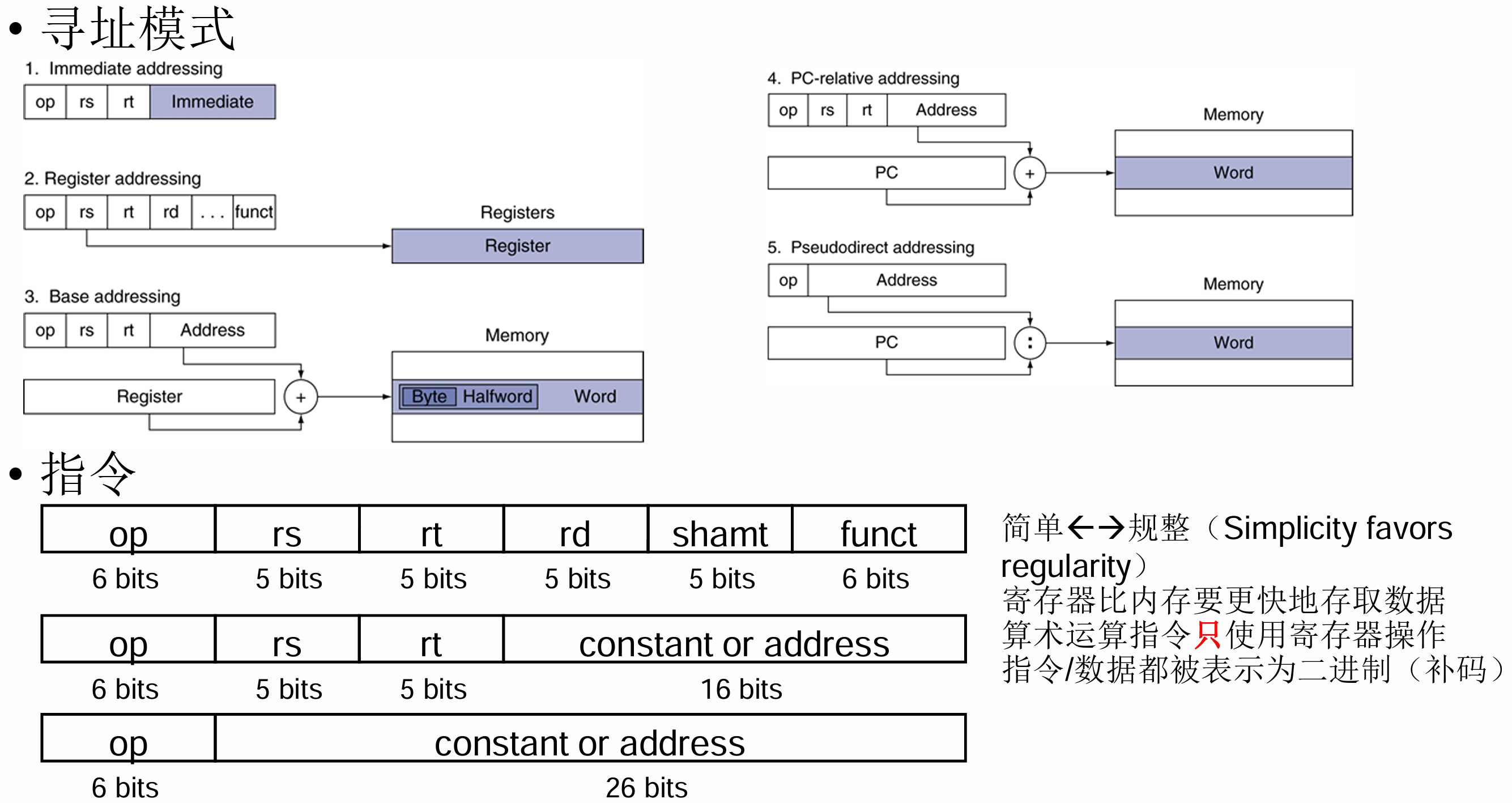
CPU时间：执行某一任务在CPU上所花费的时间，包括用户CPU 时间和系统CPU 时间

CPU时间 = CPU时钟周期数 × 时钟周期时间 = CPU时钟周期数 / 时钟频率 Tcpu = Ncpu / fclk

指令平均时钟周期：Clock Cycle per Instruction, CPI，一个指令需要的时钟周期数

CPU时钟周期数 = 指令数×CPI Ncpu = Ni × CPIi

COMS晶体管功耗=0.5×负载电容×电压²×开关频率 P=½×C×V²×F



局部性原理

时间局部性：如果某个数据项被访问，在不久的将来它可能再次被访问

空间局部性：如果某个数据项被访问，与它地址相邻的数据项可能很快也将被访问

平均存储器访问时间Average memory access time (AMAT) = 命中时间+ 缺失率×缺失代价

